DEPARTMENT OF APPLIED STATISTICS,
UNIVERSITY COLLEGE, UNIVERSITY OF LONDON

DRAPERS' COMPANY RESEARCH MEMOIRS

STUDIES IN NATIONAL DETERIORATION. VII.

THE INTENSITY OF NATURAL SELECTION IN MAN

BY

E. C. SNOW, M.A.
BIOMETRIC LABORATORY, UNIVERSITY COLLEGE, LONDON

LONDON:
PUBLISHED BY DULAU AND CO., LTD., 37, SOHO SQUARE, W.
1911

Price Three Shillings.

UNIVERSITY OF LONDON
GALTON LABORATORY OF NATIONAL EUGENICS

THE LABORATORY PUBLICATIONS ARE ISSUED BY
MESSRS DULAU AND CO., LTD., 37, SOHO SQUARE, LONDON, W.

Memoir Series.

I. **The Inheritance of Ability.** Being a statistical Examination of the Oxford Class Lists from the year 1800 onwards, and of the School Lists of Harrow and Charterhouse. By EDGAR SCHUSTER, M.A., First Galton Research Fellow in National Eugenics, and E. M. ELDERTON, Galton Research Scholar in National Eugenics. *Issued.* Price 4*s. net.*

II. **A First Study of the Statistics of Insanity and the Inheritance of the Insane** Diathesis. By DAVID HERON, M.A., Second Galton Research Fellow. *Issued.* Price 3*s. net.*

III. **The Promise of Youth and the Performance of Manhood.** Being a statistical Examination into the Relation existing between Success in the Examinations for the B.A. Degree at Oxford and subsequent Success in professional Life. (The professions considered are the Bar and the Church.) By EDGAR SCHUSTER, M.A., D.Sc., First Galton Research Fellow in National Eugenics. *Issued.* Price 2*s.* 6*d. net.*

IV. **On the Measure of the Resemblance of First Cousins.** By ETHEL M. ELDERTON, Galton Research Scholar, assisted by KARL PEARSON, F.R.S. *Issued.* Price 3*s.* 6*d. net.*

V. **A First Study of the Inheritance of Vision and of the Relative Influence** of Heredity and Environment on Sight. By AMY BARRINGTON and KARL PEARSON, F.R.S. *Issued.* Price 4*s. net.*

VI. **Treasury of Human Inheritance** (Pedigrees of physical, psychical, and pathological Characters in Man). Parts I and II (double part). (Diabetes insipidus, Split-Foot, Polydactylism, Brachydactylism, Tuberculosis, Deaf-Mutism, and Legal Ability.) Issued by the Galton Laboratory. Price 14*s. net.* Subscriptions to each set of four parts, 20*s. net.*

VII. **The Influence of Parental Occupation and Home Conditions on the** Physique of the Offspring. By ETHEL M. ELDERTON, Galton Research Scholar. *Shortly.*

VIII. **The Influence of Unfavourable Home Environment and Defective Physique** on the Intelligence of School Children. By DAVID HERON, M.A., Second Galton Research Fellow. *Issued.* Price 4*s. net.*

IX. **The Treasury of Human Inheritance** (Pedigrees of physical, psychical, and pathological Characters in Man). Part III. (Angioneurotic Oedema, Hermaphroditism, Deaf-mutism, Insanity, Commercial Ability.) Issued by the Galton Laboratory. Price 6*s. net.*

X. **A First Study of the Influence of Parental Alcoholism on the Physique** and Intelligence of the Offspring. By ETHEL M. ELDERTON, Galton Research Scholar, assisted by KARL PEARSON, F.R.S. *Issued.* Price 4*s. net.*

XI. **The Treasury of Human Inheritance** (Pedigrees of physical, psychical, and pathological Characters in Man). Part IV. (Cleft Palate, Hare-Lip, Deaf-Mutism, and Congenital Cataract.) *Issued.* Price 10*s. net.*

XII. **The Treasury of Human Inheritance** (Pedigrees of physical, psychical, and pathological Characters in Man). Parts V and VI. (Haemophilia.) *Issued.* Price 15*s. net.*

XIII. **A Second Study of the Influence of Parental Alcoholism on the Physique** and Intelligence of the Offspring. A Reply to certain Medical critics and an examination of the rebutting evidence cited by them. By KARL PEARSON, F.R.S., and ETHEL M. ELDERTON. *Issued.* Price 4*s. net.*

XIV. **A Preliminary Study of Extreme Alcoholism in Adults.** By AMY BARRINGTON and KARL PEARSON, F.R.S., assisted by DAVID HERON, D.Sc. *Issued.* Price 4*s. net.*

XV. **The Treasury of Human Inheritance** (Pedigrees of physical, psychical, and pathological Characters in Man). Parts VII and VIII. (Dwarfism.) With 50 Plates of illustrations and 8 Plates of Pedigrees. *Immediately.*

Lecture Series.

I. **The Scope and Importance to the State of the Science of National Eugenics.** By KARL PEARSON, F.R.S. *Issued.* Third Edition. Price 1*s. net.*

II. **The Groundwork of Eugenics.** By KARL PEARSON, F.R.S. *Issued.* Price 1*s. net.*

III. **The Relative Strength of Nurture and Nature.** By ETHEL M. ELDERTON. *Issued.* Price 1*s. net.*

IV. **On the Marriage of First Cousins.** By ETHEL M. ELDERTON. *Issued.* Price 1*s. net.*

V. **The Problem of Practical Eugenics.** By KARL PEARSON, F.R.S. *Issued.* Price 1*s. net.*

VI. **Nature and Nurture, the Problem of the Future.** By KARL PEARSON, F.R.S. *Issued.* Price 1*s. net.*

VII. **The Academic Aspect of the Science of National Eugenics.** By KARL PEARSON, F.R.S. *Issued.* Price 1*s. net.*

BIOMETRIC LABORATORY, UNIVERSITY COLLEGE, LONDON

DRAPERS' COMPANY RESEARCH MEMOIRS

STUDIES IN NATIONAL DETERIORATION. VII.

THE INTENSITY OF NATURAL SELECTION IN MAN

BY

E. C. SNOW, M.A.
BIOMETRIC LABORATORY, UNIVERSITY COLLEGE, LONDON

LONDON:
PUBLISHED BY DULAU AND CO., LTD., 37, SOHO SQUARE, W.
1911

PREFATORY NOTE.

In publishing Mr Snow's memoir on the intensity of natural selection in man I have been influenced by certain facts which it is desirable that the reader also should bear in mind. The chief problem that impresses itself upon all social inquirers at the present day centres in the question of whether the enormous growth of charitable institutions and the municipal provision made for the poor and destitute, unaccompanied as they now are by any effective limitation on reproduction, are not tending towards the degeneration of the race by handicapping the fitter for the sake of the less fit members of society. The problem is undoubtedly a most difficult one. It can only be attacked on purely academic lines; no real solution will be obtained unless it is studied in a manner wholly free of party bias and by those having no responsibility either to the minister of a political party, or to a government department.

The Local Government Board has recently seen fit to issue a series of Reports, which deal with some of the most difficult statistical problems, and draw definite conclusions on very insufficient data by wholly inadequate methods. Nearly every one of these Reports has reached results which, if correct are certainly unproven, and which are often demonstrably incorrect. The recent Report on Infant and Child Mortality is a striking instance of this kind. The problem apparently set to the writer was precisely this problem of selection; and although the writer does not definitely state that no such selection exists, he indulges in such ambiguous phrases that the newspaper press has almost universally assumed that a heavy child mortality does not leave a relatively stronger population, and *pro tanto* argued that Darwinism has little application to mankind. The data the author has dealt with could not as they stand possibly answer the problem set to him. The problem is an extremely difficult and complex one abounding in most dangerous pitfalls, and light can be thrown upon it only by lengthy and arduous research. No one will blame a very busy government official for having neither the leisure nor the training for such an investigation. But the problem of whether Darwinism does or does not apply to man is essentially an academic problem, and not one which can be settled in a government report, following a minister's instructions to a subordinate to investigate it. Mr Snow's work may or may not be final, but it shows at once the extreme difficulties in the path of the scientific investigator in this field, and indicates how little the writer of the report in question has realised those difficulties. No one on the basis of a scientific memoir is likely to propose sweeping legislation; such a memoir when issued meets the criticism of trained scientists, and eventually either becomes part of the common stock of knowledge, or its results are modified and reconstructed. But a government publication professing, even ambiguously, to answer a great scientific problem, is a real social danger; its views are read and accepted in political circles, where no scientific criticism is available, and legislation may follow on the basis of a report, which differs even from that of a Royal Commission, in that, while it is a one-man production, it is nevertheless backed by all the weight which attaches to an official government document.

KARL PEARSON.

The Intensity of Natural Selection in Man.

By E. C. SNOW, M.A., *Biometric Laboratory, University College, London.*

> "Every man who has lived through a hard winter, every man who has examined a mortality table, every man who has studied the history of nations has probably seen natural selection at work. It is not the existence of natural selection, but its *intensity* and *rapidity* of action in the transmutation of species, which is the problem...." K. Pearson (1894), *The Chances of Death*, Vol. I. p. 166.

(I) The object of the present memoir lies in an attempt to measure the rate at which natural selection, as shown by a selective death-rate, is operating in a general human population. No effort will be made to answer the question *how*, if at all, does natural selection produce its effects in man. There is as yet little reliable data available which can permit of a satisfactory reply being given to such a query. We cannot at present say with certainty, for example, whether the fair featured people are killed off faster than the general population, nor whether the brachycephalic have more chance of survival, although there is some evidence for these views. The data accruing from the compulsory medical inspection of school children should soon be sufficient to render possible an answer to these and many similar questions in the case of man which have already been partially answered by Weldon, Harris and others for lower forms of life. It is quite possible, however, to enquire into the problem: What proportion of the death-rate in a general population is selective?

The data used will be gathered from the widest available sources—national mortality statistics. From the point of view of evolution such data is of more value than that which can be gathered from genealogies and family records. The latter, though frequently more complete, refer only to a restricted and particular class of society and not to a mass population. But in order to show the evolutionary importance of natural selection its existence must be demonstrated in a perfectly general population.

A complete account of the work performed will be given, though the data for some of this was felt to be unsatisfactory and some contradictory results were obtained. In the opinion of the author those results which point most clearly to the action of selection were derived from the more reliable data, but to bring out the difficulties that were met with and for comparison with other recent work on the subject, to which reference will be made, it seemed desirable to omit no part of the investigation. It is more in accordance with scientific spirit that the reader should be allowed to draw his conclusions from the whole of the research, and to form his own opinion on the value of the material used and of the results deduced from it. Those who hold a particular view must then show that the results which most favour that view are founded on the better data.

(II) The question of the application of the principle of natural selection to man is one of great importance, both to scientists and to sociologists. There are still a few of the former, as there were fifty years ago, who believe that although natural selection applies to the rest of animal and plant life it does not so apply to man under his present conditions of existence. Yet there is far more material available to test the truth of the Darwinian hypothesis in the case of man than in the case of any other organism, though the methods by which this can be adequately performed have only been developed in the last few years.

Darwin, who demonstrated clearly the action of the survival of the fittest in primitive societies, brought forth no conclusive evidence for its operation in modern civilised communities. It was as a working hypothesis to be applied to man, and not as a demonstrated fact, that the Darwinian principle had the approval of Mill as logically sound. The weight of Darwin's arguments when applied to lower forms of life depended almost entirely on the restriction that the population dealt with was practically stationary. Many quotations from Darwin and his commentators could be made to emphasize this point*. An apposite one from Huxley may be given: "Year after year, the wild animals with which man never interferes are, on the average, neither more nor less numerous than they were; and yet we know that the annual produce of every pair is from one to perhaps a million young; so that it is mathematically certain that, on the average, as many are killed by natural causes as are born every year, and those only escape which happen to be a little better fitted to resist destruction than those which die." But we cannot legitimately extend the argument to civilised man. The population of most civilised countries is increasing continuously, and not a very large proportion perishes before the reproductive age is reached. A reference to any "life-table" will confirm this. More scientifically, Pearson by analysing a mortality curve found that it could be split up into five "chance" distributions, each of which represents a component mortality curve centring about a particular age. For 1000 English males born in a particular year he found that their ages at death could be summarised thus:

Mortality of old age centring about 67,	484	deaths
Mortality of middle life centring about 41,	173	,,
Mortality of youth centring about 22,	51	,,
Mortality of childhood centring about 6,	46	,,
Mortality of infancy (mostly in the first two years)	246	,,

and he concluded that "at most a minority, perhaps 40 in the 100 males, die before reaching an age at which they could many times have reproduced themselves." What proportion of this 40 °/o who die before their reproductive activity is completed is selected?

---

* See Pearson's *The Chances of Death*, I. p. 123.

Information as to the rate at which civilised communities die could have been easily made out from the national mortality statistics when the evolutionary controversy was at its height. Yet references to the Registrar-General's statistics—containing valuable data even before the *Origin of Species* was published—are extremely meagre in Darwinian literature. The possibility of bringing the resources of Somerset House or any other store-house of statistics to the aid of science seems to have been overlooked until the rise of the Biometric School and the introduction of exact quantitative methods towards the end of last century. The measurement of natural selection was one of the first problems to which its attention was turned. In a number of papers, Pearson and his collaborators did much to ascertain the intensity of selection in man under modern conditions, and in other forms of life. Judged by recent statements in scientific and pseudo-scientific journals Pearson's work seems to have been overlooked, and any evidence, however slight or untrustworthy, which has appeared to deny the operation of natural selection in man has been accepted with little question. It will be well, therefore, to reproduce here the arguments used by Pearson.

(III) In the first place, distinction must be made between selection arising from a differential death-rate and that arising from a differential birth-rate. The latter—selection arising through differences of fertility—Pearson has termed "reproductive selection." Its existence in civilised communities at the present time is undoubted, and has already received the attention of legislators. Much data has been collected in regard to it in England, and considerably more will be available when the results of the recent census are disclosed. But it is that form of selection which is due to differences in mortality with which we wish to deal in this paper. The effects of this are probably almost completed before the reproductive age.

Pearson's demonstration of the existence of a selective death-rate is not difficult to follow. In his first paper on the subject* the correlations between the length of life of father and son, and also of brothers, were worked out. The restriction to the male line was enforced by the limitations of the data, which were gathered from Foster's *Peerage* and Burke's *Landed Gentry*. This restriction was removed for the second paper†, and more copious material was obtained from extracts made from the genealogical records of the Society of Friends. From this many more correlations were worked out. In direct inheritance those for the duration of life between father and both adult and minor sons and daughters were found, while corresponding results were obtained for mother and the same offspring. In collateral inheritance, the correlations for duration of life between all possible combinations of pairs of adult and minor brothers and sisters were ascertained. Between parents and adult children the average correlation for all combinations of sexes was found to be ·136, while for adult siblings it was ·283.

---

* "Data for Problem of Evolution in Man, II. A First Study of the Inheritance of Longevity and the Selective Death-rate in Man." By Mary Beeton and Karl Pearson, F.R.S. *Proc. R. S.* Vol. 65, pp. 290—305.

† *Biometrika*, Vol. I. Oct. 1901, pp. 50—89.

At the time these papers were written, a number of investigations had been made on a variety of characters and organs in man, and average values of the correlations found were ·4 between parent and offspring and ·45 between siblings. (More recent work has placed these correlations higher.) Thus the intensity of correlation for duration of life is much lower than for other inheritable characters. Pearson considered that the true explanation of this lay in the fact that the death-rate is really of two kinds, one a non-selective (purely random) rate—by which the weak and strong are stricken alike—and the other wholly selective. The arrows of the latter kind of death, though striking at all, kill only those of weakly constitution. The consequence of this is that in forming a correlation table for ages at death of parent and offspring, or of pairs of siblings, we are dealing with a mixture of material, viz. of those deaths which are due to purely random causes and those due to selective causes. Were we dealing with the latter kind alone, we should expect a correlation of approximately equal value to that found for other characters or organs, while if we were dealing only with the random causes of death no relationship at all should be expected between ages at death of father and son, or of siblings. Now, when correlated material is mixed with uncorrelated material, it has been shown* that the result is approximately to reduce the coefficient of correlation in the ratio of the amount of the correlated to the total amount of material. By this theorem, then, we are enabled to determine what proportion of the total number of deaths dealt with in forming the correlation table is due to selective action. Taking values of the parental correlation from ·3 to ·45, Pearson computed that from 55 °/。 to 67 °/。 of the total deaths in the material he dealt with were ascribable to selective action. In the same way, taking values of the fraternal correlation between ·4 and ·5 his results showed from 75 °/。 to 84 °/。 of the deaths to be due to selection. If we take the mean of the many values found by more recent research for the parental and fraternal correlations†—·470 and ·521 respectively—the proportions of the total deaths which were effected by selective action work out at 53·9 and 73·7 °/。 in the two cases. The divergence between the results derived from fraternal and parental correlation was attributed to the environment changing much more from parent to offspring than it does from brother to brother. Environmental conditions are usually more alike for sisters than for brothers, and Pearson found that the correlation between length of life of sisters was as high as ·3322. This value would lead to an even higher proportion of selected deaths.

The crude relationship between length of life of parent and offspring can be seen very well from the following tables deduced from Pearson's data‡.

---

* Pearson, *Phil. Trans.* A, Vol. 192, p. 277.

† References to these will be found in *Proc. R. S.* B, Vol. 83, 1910, p. 54.

‡ "Groundwork of Eugenics," *Eugenics Laboratory Lecture Series II*, p. 25.

A. Mothers and Daughters.

Age of Mother at Death.

Offspring	19—29	29—39	39—49	49—59	59—69	69—79	79—89	89 and over	Totals
Died before 21	36	122	123	111	115	162	91	22	782
Survived 21 ...	13	63	93	77	180	291	283	64	1064
Totals	49	185	216	188	295	453	374	86	1846
Death-rate......	73·5	65·9	56·9	59·0	39·0	35·8	24·3	25·6	42·4

B. Fathers and Daughters.

Age of Father at Death.

Offspring	24—34	34—44	44—54	54—64	64—74	74—84	84—94	94 and over	Totals
Died before 21	34	83	118	144	202	194	71	7	853
Survived 21 ...	13	43	100	168	309	365	142	16	1156
Totals	47	126	218	312	511	559	213	23	2009
Death-rate......	72·3	65·9	54·1	46·2	39·5	34·7	33·3	30·4	42·5

Thus 68 °/o of the children born of mothers who died before the age of 39 themselves died before 21, while of fathers who died before the age of 44 an equal proportion of offspring died as minors. But not much more than 30 °/o of the children born of parents who reached the 70th year died before they were 21. These facts point just as strongly to the action of selection as do the correlation coefficients given above.

Prof. Ploetz of Munich[*] has recently confirmed Pearson's conclusions. Working on two series of data, one supplied by middle-class and the other by royal families, he found that at least 60 °/o of the mortality at ages less than five was due to selective causes.

Results obtained from such restricted data as that described above cannot, of course, be considered as fairly representative of the general population. The social status of the members of the Society of Friends is considerably above that of the general population, but there is no *a priori* reason why selection should depend directly on social standing. On the contrary, the masses of the population live under far more exacting conditions, and members of weak stocks among them have no chance of migration to more congenial climates in order that their lives may be prolonged. There seems no cause for the view, then, that the results of Pearson and Ploetz were at all weighted by reason of the social standing of the classes dealt with[†].

---

[*] *Archiv für Rassen- und Gesellschafts-Biologie*, Bd. VI. S. 33, 1909.

[†] The correlation will be lessened rather than magnified by taking a single group out of the whole population; hence Pearson's percentages of selective deaths are likely to be minimum values.

Another possible point of criticism lies in the fact that frequently in genealogies many of the deaths of minors are unrecorded. This was frankly admitted by those working on the material supplied by the records of the Society of Friends, and a number of conclusions which might have been made were omitted on account of doubt on this point. In order to affect the results found for the intensity of selection, however, it is necessary to demonstrate that the omission of the deaths of minors was differential and depended upon the age at death of parent. There is no more reason to believe that the death of a minor should go unrecorded in the genealogies when the parents die early than when they die late, and we are of the opinion, therefore, that it is quite justifiable to make the inference that natural selection is strongly operative in the general population. Nevertheless, a direct demonstration of its operation and a measure of its intensity in the general population would be extremely desirable and valuable.

(IV) An attempt in this direction has lately been made in a Supplement to the Report of the Medical Officer of the Local Government Board (Dr Newsholme) on Infant and Child Mortality. This report marks an important stage in official statistical publications, inasmuch as it is the first of such publications in which the subject has been investigated by the modern method of correlation, by which quantitative measures of the various factors of relationship can be made out. It contains also a suggestive appendix by Mr Udny Yule "On the Possible Selective Influence of Mortality in Infancy on Mortality in the Next Four Years of Life." He concluded "that there is little definite evidence of such selection beyond the second year of life, and that after the third year the environmental influences even of infancy alone appear to preponderate over any possible selective influence." Mr Yule, however, added the proviso that "the statistical basis is too slight, and the case too complex to enable one to put forward such a conclusion on the basis of these figures alone as anything but a very tentative hypothesis," a provision which is overlooked in the body of the report. On account of the apparent contradiction of these results with those previously reached by Pearson, and the general importance of the subject, a fuller investigation by means of our national mortality statistics seemed to be called for. This is the more necessary as it appeared to the present author that the data supplied to Mr Yule were inadequate and unsatisfactory for such an important enquiry, while considerable commendatory attention has been given to the results deduced from it both in the daily and in the scientific press.

A little space may here be given to the results and conclusions of the report. One of its three objects was "to determine, on the basis of our national statistics, whether reduction of infant mortality implies any untoward influence on the health of survivors to later years." By this statement it appears that the investigation was specifically intended to discover whether selection exists under modern conditions. Part I was explicitly directed to this end and the point of view of selection combated. "...Attempts to reduce infant mortality are regarded "by many as an interference with natural selection, which must be inimical to the "average health of those surviving. According to this school of thought efforts to "save infant life merely prevent the 'weeding out' of the unfit, and ensure the "survival of an excessive proportion of weaklings. The statistics in the following "pages do not support this view, for, as shown in Table I of Appendix III, "counties having a high infant mortality have also a high death-rate at ages 1—5 "and this higher death-rate will be seen shortly to continue at higher ages. The "converse holds good for the counties having a low infant mortality." The table referred to is headed "1908—Deaths of Infants under 1 year of age per 1000 Births, and at 1—5 per 1000 Survivors at age of one, in the Counties of England and Wales," and it brings out clearly a fact which is well emphasised by a number of diagrams and frequently reiterated that, on the whole, a county which had a high or low infant mortality in 1908 had also a correspondingly high or low mortality for ages 1—5 *in that same year.* In addition the correlation between the infantile death-rate and that for children between 1 and 5 for 43 counties in the year 1908 was calculated and found to be ·88 ± ·023. These figures, however, have really no bearing whatever on the question of selection. They bring out in a very emphatic manner a continuity in the environmental conditions operating on all between ages 0 and 5 alike. They do not deal at all with "the average health of those surviving" the infant period considered. To demonstrate effectively the non-action of natural selection some other means must be adopted, and that "school of thought" which has based its belief of the prevalence of selection on the facts established by Pearson and Ploetz need have no fear that Part I of the report will overthrow its creed.

Later in the report (p. 17), in contradiction to the paragraph quoted above, it is averred that "no attempt has been made in the preceding pages to settle the most important point as to whether a heavy infant mortality has any selective influence on the population surviving beyond infancy." The reader of the "prefatory remarks," however, must certainly conclude that Part I was intended to deal with the question of selection, and Part II opens by repeating the ambiguous phrase "an excessive infant mortality is nearly always associated with an excessive death-rate during the next four years of life." The term "next four years of life" implies that the particular population dealt with has been followed through the first five years of its existence and mortality figures collected for each of these years. This is the interpretation apparently adopted by a number of the reviewers of the report. It cannot be too strongly emphasised that all that has been demonstrated in Part I is that *in the same year* infant mortality and mortality at higher ages in childhood are positively correlated. The intended object of Part I—the consideration of the health of *survivors* to later years—has not been attempted there.

That matter is, however, considered by Mr Yule in his appendix, and this is the only part of the report in which natural selection (beyond its mere definition) is really dealt with. As before stated we think the data supplied to Mr Yule quite insufficient to discover any selective influences in a general population. The appendix exhibits considerable compression and it is not possible to give a short account of Mr Yule's work. He used the figures for the whole of England and

Wales from 1855 to 1908. Thus he followed those born in 1855 through their first five years of life to 1859; those born in 1856 he followed through to 1860, and so on, stopping with those born in 1904. For full details of the work we must refer to the appendix itself; the conclusions and caution have previously been stated.

The objection to the data used by Mr Yule is that it referred to the whole of England and Wales and not to its constituent parts separately. Much is lost through this. For in a particular locality there is considerable variation in the infantile and child death-rates from year to year owing to *local* epidemics caused, *e.g.*, by meteorological conditions. Such epidemics do not occur in every district in the same year nor, in general, in the same district in succeeding years. The weather conditions in the south-west of England may, in a particular year, cause the infant mortality there to be much above its average for the previous few years. In the north-east, however, in the same year the conditions may be much more favourable, and the infant mortality below its previous average. From the point of view of selection we should wish to compare the after history of those born in that particular year in the former area with those born in the same year in the latter, correction being made to render the after environment of the two populations the same. But in clubbing the districts together we are losing the advantage of these local variations. If, for the sake of argument, we suppose selection were operative, and in the first year one half of the country from random causes had high mortality and the other half low (environmental conditions being the same), while in a subsequent year or series of years the first half had low and the second high mortality, this selection would be lost sight of altogether when the districts were grouped into one whole.

That this is not a mere theoretical criticism can be seen from the tables given at the end of this memoir. Taking the six areas at the top of Table IV and the six at the bottom—the former covering much of the rural region in the south of England, and the latter in the north—we find that the number of males who were born in 1871 and who died in the first three years of life were 3774 and 4205 respectively. But for those born in 1872 the corresponding figures were 3580 and 4343, the populations on which these figures were based not altering appreciably in the periods. Thus while mortality in the south diminished that in the north increased, but by clubbing the areas together the effects of these opposite movements are nullified. The schedules from which the table was constructed permit of the corresponding figures for infant deaths only being given. We find that in the northern areas the mortality of male infants increased from 2986 per 20,563 births in 1871 to 3194 per 20,963 births in 1872, but in the southern areas there was a diminution from 2759 per 20,277 births to 2655 per 20,834 births. In taking the whole country then we are dealing with the resultant of many tendencies which are of all sorts of magnitudes. Though this resultant does not vary much from year to year the individual components may alter greatly. It is essential in this as in many other problems that we take our units sufficiently small so that the effects we are investigating may be considered fairly uniform and homogeneous in the area*.

---

* For the numerical importance of this: see p. 20 (iv) below.

(V) The problem to be investigated may be stated thus: Consider the infants born in a number of districts and in a particular year, say 1903. Follow these individual cohorts of children throughout life as far as statistics permit. Allowing for differences of environment, so far as this can be done, will the cohort which had a high mortality in its first year or two of life have, on the whole, a high or low mortality in the subsequent few years of its existence? To do this we shall want to know,

the deaths under 1 year of each cohort in 1903,
,, ,, between 1 and 2 years ,, ,, ,, ,, 1904,
,, ,, ,, 2 and 3 ,, ,, ,, ,, ,, 1905,

and so on. We can then correlate the deaths, say, in the first three years of life with those in the 4th and 5th together for as many districts as possible, correction being made for constant size of cohort and for constant environment. Three points of difficulty arise here. In the first place, a small proportion of the deaths of those less than one year in the particular year, 1903, will not be of those children born in that same year, but in the previous one, and similarly for the deaths in subsequent years. This, however, is not likely to affect the argument when a large number of districts is taken, especially as over one half of the infant deaths occur in the first three months of life (over one-third in the first month). It would affect the argument if there were great seasonal fluctuations in the birth-rate, or if environmental conditions suddenly changed at the end of a year. These conditions do not occur in England. Some unit of time must be taken, and the year is the only one possible with the statistics available.

In the second place, when we have corrected for constant environment and constant size of cohort, we should expect a negative result for the above correlation, even if selection were inoperative. For in a district in which the number of deaths in the first period was above the average (this arising merely from random causes) the available population at which death could strike in the second period would be below the average, and consequently the most probable number of deaths in this district in the second period would also be below the average. Conversely, a district in which the number of deaths in the first period was below the average would from random causes alone have its deaths above the average in the second. As we are dealing only with the first few years of life *a priori* considerations would lead us to expect that this negative correlation would be small, and some idea of its magnitude will be given later (§ XXIV). But we can agree that a small negative correlation, even if significant with regard to its probable error, does not necessarily denote selection, a constant environment being assumed.

The third source of difficulty lies in the interpretation of a "constant environment." No general measure of environment seems possible: that for a particular problem must be determined by the conditions of that problem alone. A general discussion of the measurement of environment need not therefore be entered upon; the subject will be considered in each case as the necessity arises.

(VI) The data for the first attempt at the problem were taken from the various Annual Reports of the Registrar-General for England and Wales. The deaths in each year of life up to the fifth are given for registration counties in the recent volumes. The births in those areas are given elsewhere in the same volumes. The year 1903 was chosen, and the children born in that year followed through the first five years of their lives. It was recognised from the beginning that the data were not altogether satisfactory on account of the extreme differences in size and heterogeneity in conditions of the registration counties. These remained even when Huntingdonshire was grouped with Bedfordshire, Rutland with Leicester, and the whole of North Wales taken together. But as no more satisfactory material was available for recent years the work was continued. 42 districts in all were used, but in some of the work the eight largest—London, Essex, Staffordshire, Warwickshire, Lancashire, West Riding of Yorkshire, Durham and Glamorganshire,—were omitted, and the 34 of more moderate size used. In one case, too, the figures were weighted according to the sizes of the districts—as shown roughly by the number of births. To have got the areas of approximate equality in the numbers of births would have involved clubbing together so many of the smaller ones that there would have been only ten units to deal with. The crude data from which the correlations were obtained are collected in Table I. The schedules contained considerably more which could have been employed.

The measure of environment which was used in this case was the death-rate of children under five during the *preceding* decade 1890—1900. Clearly the conditions of the problem render it advisable to use only the deaths under five, and those in the decade mentioned were taken because they were readily accessible from the last Decennial Supplement of the Registrar-General. This would probably not be a bad measure of environment if the latter had remained fairly permanent in the individual districts up to the period dealt with, or had changed uniformly throughout the districts. But during the present century the death-rate of children under five has been falling, but not uniformly throughout the districts. The reduction has been greatest in the urban districts, which, however, still have the highest child mortality. The measure of environment, therefore, cannot be considered a very good one.

It was at first intended to correlate the deaths in the first year of life with deaths in the second for children born in 1903, for a constant number of births and a constant environment. But even if selection acted stringently during the first two years of life its existence might not be brought out by such a correlation. For there is no clear line of demarcation between the conditions of the first year and those of the second. Pearson found that for true infantile mortality 53 deaths occur in the second year for every 156 which occur in the first. It accordingly seemed best for a start to correlate deaths in the first year of life with those in the fifth. In this way we should see if a high infantile mortality left a population more immune in its fifth year of existence. A possible defect in this is that selection may not have had its full innings by the close of the first year, and that a longer period should therefore be used. This is subsequently done.

Using the following notation :

Births in 1903 .... x_0,
Deaths under 1 year in 1903 .... x_1,
Deaths between 4 and 5 in 1907 .... x_2,
Total Deaths between 0 and 5, 1890—1900 ...... x_3,
Mean population between 0 and 5, 1890—1900... x_4,
Death-rate under 5 $\left(i.e.\ 100\,.\dfrac{x_3}{x_4}\right)$, 1890—1900 ... x_5,

we find for the 42 districts the following constants* :

Means	Standard Deviations	Correlations	
$\bar{x}_0 = 22600$	$\sigma_0 = 28150$	$r_{01} = \cdot988$	$r_{40} = \cdot995$
$\bar{x}_1 = 2970$	$\sigma_1 = 4160$	$r_{02} = \cdot992$	$r_{41} = \cdot995$
$\bar{x}_2 = 133$	$\sigma_2 = 189$	$r_{03} = \cdot998$	$r_{42} = \cdot992$
$\bar{x}_3 = 49980$	$\sigma_3 = 75060$	$r_{12} = \cdot994$	$r_{43} = \cdot996$
$\bar{x}_4 = 86520$	$\sigma_4 = 104640$	$r_{13} = \cdot990$	$r_{50} = \cdot643$
$\bar{x}_5 = 49\cdot4$	$\sigma_5 = 13\cdot6$	$r_{23} = \cdot987$	$r_{51} = \cdot676$
			$r_{52} = \cdot669$

The crude correlation between the deaths in the first year and deaths in the fifth was ·994. When corrected for a constant number of births this was reduced to ·571 (${}_0r_{12}$). When in addition it was corrected for a constant environment rate it was still further reduced to ·500 (${}_{05}r_{18}$). The correction for absolute environment, however, took the correlation up to ·591 (${}_{034}r_{12}$).

The correlations obtained by using death-*rates* instead of absolute numbers of deaths were also worked out. If

$x_6 =$ Infant Death-rate (per 1000 births) in 1903,
$x_7 =$ 5 year old Death-rate (per 1000 survivors to that age) in 1907,

then $\bar{x}_6 = 118\cdot14,\quad \sigma_6 = 21\cdot91,\quad r_{67} = \cdot782,$
$\bar{x}_7 = 6\cdot33,\quad \sigma_7 = 1\cdot70,\quad r_{56} = \cdot715,\quad r_{57} = \cdot592.$

From these we find that the correlation between infant mortality and 5 year old mortality of those born in 1903 for a constant environment rate was ·636 (${}_5r_{67}$).

The corresponding correlations were worked out when the observations were weighted according to the number of births in the district. The constants then were :

$$\bar{x}_6 = 145\cdot2,\quad \sigma_6 = 91\cdot8,\quad r_{67} = \cdot974,$$
$$\bar{x}_7 = 8\cdot3,\quad \sigma_7 = 67\cdot9,\quad r_{57} = \cdot934,\quad r_{56} = \cdot937,$$
$$\bar{x}_5 = 68\cdot0,\quad \sigma_5 = 64\cdot0.$$

In this case we find ${}_5r_{67} = \cdot793$.

These figures give no support whatever to the supposed operation of selection.

---

* The correlations were actually worked out to eight places of decimals. This is very necessary when the correlations are so high, as an omission in an ordinary correlation may be magnified hundreds of times when a partial correlation is based on it.

(VII) The eight largest districts were next omitted and correlations worked out. These are now smaller, and are therefore more satisfactory to base a partial correlation on.

Using the same notation as in the last paragraph, the statistical constants are:

Means	Standard Deviations	Correlations	
$\bar{x}_0 = 12420$	$\sigma_0 = 5810$	$r_{01} = \cdot 949$	$r_{23} = \cdot 846$
$\bar{x}_1 = 1450$	$\sigma_1 = 790$	$r_{02} = \cdot 918$	$r_{40} = \cdot 954$
$\bar{x}_2 = 64$	$\sigma_2 = 39$	$r_{03} = \cdot 971$	$r_{41} = \cdot 971$
$\bar{x}_3 = 23570$	$\sigma_3 = 11262$	$r_{12} = \cdot 958$	$r_{42} = \cdot 913$
$\bar{x}_4 = 49930$	$\sigma_4 = 20100$	$r_{13} = \cdot 891$	$r_{43} = \cdot 943$

The partial correlations are:

$$_0r_{12} = \cdot 696, \quad {}_{034}r_{12} = \cdot 456.$$

This last figure is smaller than that obtained for the whole 42 districts, but still indicates that there is a significant permanence in high or low mortality between the first and fifth years of life, even though correction has been attempted for environment.

(VIII) An alteration in the procedure was next made. The first three years of life was the period taken for selection to act, and the effects of differences of mortality in this period on the mortality in the fourth and fifth years of life conjointly investigated. For another case, the first four years of life was taken for the one period, and the fifth for the other. An altogether different measure of environment, too, was attempted. This was to make constant in each district all the deaths under five years of age in the five years 1903—1907 other than the deaths of the particular population (viz. those born in 1903) dealt with. The actual size of the population on which these other deaths are based is rendered approximately the same throughout the districts by correcting for a constant number of births. To obtain the exact population on which these "other deaths" in the first five years of life are based is a long and laborious work, but it is attempted in a particular case later on. For the present we shall assume that fixing the number of births in a particular year does fairly closely fix the size of the population under five years of age in the next five years.

These "other deaths" under five years of age will be about four-fifths of the total deaths under that age in the five years considered. A better measure of environment would be to fix the total deaths under five (the population on which these are based being constant). But this would include the deaths used in the other variables and a possibility of spurious correlation would arise. Making constant only these four-fifths of the total deaths will mean that we do not sufficiently correct for environment. Instead of having the environment varying between very wide limits it will now only vary between very narrow limits, but there will still be districts in which the environment is a little worse than for others. Thus from such conditions alone there will be a tendency towards a positive correlation between the deaths in the two periods. If therefore we find a negative correlation when corrected for this approximate environment, this will be less than that which would be reached if the measure of environment were perfect.

(IX) For the three-year and two-year periods the variables then are:

x_0 = Births in 1903.

x_1 = Deaths in first three years of life of those born in 1903.

x_2 = Deaths in fourth and fifth years of life of those born in 1903.

x_3 = The *remaining* deaths under 5 years in the years 1903—7, *i.e.* of those 1—5 in 1903, 0—1 and 2—5 in 1904, etc.

The correlations were worked out for the 34 districts and also for the whole 42. For the former the constants are:

$$\begin{array}{llll} \bar{x}_0 = 12420, & \sigma_0 = 5813, & {}^*r_{01} = \cdot 946822, & r_{12} = \cdot 975283, \\ \bar{x}_1 = 1940, & \sigma_1 = 1080, & r_{02} = \cdot 936851, & r_{13} = \cdot 996246, \\ \bar{x}_2 = 143, & \sigma_2 = 84, & r_{03} = \cdot 958442, & r_{23} = \cdot 981327, \\ \bar{x}_3 = 8020, & \sigma_3 = 4420. & & \end{array}$$

From these we find the partial correlations:

$$\begin{array}{ll} {}_3r_{01} = -\cdot 324880, & {}_3r_{02} = -\cdot 067338, \\ {}_3r_{12} = -\cdot 141869, & {}_{03}r_{12} = -\cdot 173523 \pm \cdot 1122. \end{array}$$

For the whole 42 districts the corresponding figures are:

$$\begin{array}{llll} \bar{x}_0 = 22578, & \sigma_0 = 28152, & r_{01} = \cdot 992404, & r_{12} = \cdot 998367, \\ \bar{x}_1 = 4113, & \sigma_1 = 5989, & r_{02} = \cdot 994491, & r_{13} = \cdot 999622, \\ \bar{x}_2 = 305, & \sigma_2 = 449, & r_{03} = \cdot 993749, & r_{23} = \cdot 998228, \\ \bar{x}_3 = 16921, & \sigma_3 = 24351. & & \end{array}$$

$$\begin{array}{ll} {}_3r_{01} = -\cdot 316009, & {}_3r_{02} = +\cdot 376779, \\ {}_3r_{12} = +\cdot 315985, & {}_{03}r_{12} = +\cdot 495030. \end{array}$$

(X) For the four-year and one-year periods we have the extra variables:

x_4 = Deaths in first four years of life of those born in 1903.

x_5 = Deaths in fifth year of life of those born in 1903.

x_0 and x_3 being as before.

For the 34 districts we have the additional constants:

$$\begin{array}{llll} \bar{x}_4 = 2019, & \sigma_4 = 1125, & r_{03} = \cdot 958442, & r_{34} = \cdot 996518, \\ \bar{x}_5 = 64, & \sigma_5 = 39, & r_{04} = \cdot 947308, & r_{35} = \cdot 966515, \\ & & r_{05} = \cdot 918133, & r_{45} = \cdot 957172. \end{array}$$

$$\begin{array}{ll} {}_3r_{04} = -\cdot 328043, & {}_3r_{05} = -\cdot 112214, \\ {}_3r_{45} = -\cdot 279621, & {}_{03}r_{45} = -\cdot 337107 \pm \cdot 1025. \end{array}$$

---

* These correlations were taken to eight places of decimals.

For the 42 districts we find :

$$\bar{x}_4 = 4284, \quad \sigma_4 = 6249, \quad r_{03} = {\cdot}993749, \quad r_{34} = {\cdot}999611,$$
$$\bar{x}_5 = 133, \quad \sigma_5 = 189, \quad r_{04} = {\cdot}992599, \quad r_{35} = {\cdot}998302,$$
$$r_{05} = {\cdot}992416, \quad r_{45} = {\cdot}997749.$$
$${}_3r_{04} = -{\cdot}245032, \quad {}_3r_{05} = +{\cdot}054276,$$
$${}_3r_{45} = -{\cdot}100925, \quad {}_{03}r_{45} = -{\cdot}090515.$$

(XI) In the consideration of the results given in the last two paragraphs we think that much greater weight is to be attached to those reached when the very large districts are omitted. These very large districts exercise such a predominating effect that only 7 or 8 of the whole 42 observations lie above the means of the different variables. Many of the correlations approach unity, and in order to obtain the partial correlations with accuracy they had to be taken to eight places of decimals. It is idle to pretend that, for the whole 42 districts, we are dealing with variation in homogeneous material, and the results, therefore, will receive no comment. The material for the 34 districts alone, however, though far from perfect, is much more defensible. In respect of each variable, about one-half of the observations lie on either side of the corresponding means. For the births five observations on each side lie without the belt formed by measuring the standard deviation on each side of the mean, and 24 observations within that belt. Similar relationships hold for the other variables. Without these eight large districts, the data appear to be much more homogeneous as regards general conditions, and we have no doubt that greater weight should be attached to the results obtained when they are omitted. We note that both ${}_{03}r_{12}$ and ${}_{03}r_{45}$ are negative, the former not significantly so, but the latter is more than three times its probable error. Thus when correction is made for environment, those local populations which showed a high mortality in their first three or four years of life had, on the whole, a low mortality in the remaining years up to the fifth. It appears then that selection was operative.

A single correlation, however, is not sufficiently trustworthy, even when more than three times its probable error, to draw important conclusions from. Some further investigation was desirable. This might have been performed on similar data taken from a different year. But a great advantage would be gained by dealing with districts of more uniform size and, if possible, of more homogeneous conditions. A single registration county contains almost always urban and rural portions of varying size. Buckinghamshire is almost entirely rural, Norfolk is mostly rural but contains a large industrial population in Norwich. Sussex is rural, but with a number of seaside towns in which the conditions are not purely industrial. The annual reports of the Registrar-General since 1881 do not afford any material for splitting up the counties into more homogeneous areas, but the earlier ones contain much more that is of value. In fact, from a number of points of view, the earlier reports of the Registrar-General are of far more value to the purely scientific investigator than are the later ones. In the former the deaths in the first five years of life are given both for males and females in each of the 600 odd registration districts. The large majority of these, of course, were far too small to be used separately, but they could be grouped together so that units of almost any desired size could be obtained.

(XII) The year 1870 was chosen, and those born then were followed through the first five years of their existence. This was anterior to the date which is generally agreed upon as that when artificial interference with the birth-rate—producing in itself a kind of selection—commenced. Contiguous districts were grouped together wherever necessary, so that the male births in the group in the year 1870 were between 2000 and 5000. Sometimes as many as ten registration districts were taken to form a group. The manner of arrangement can be made out from Table II. In the ensuing pages such groups of districts will be referred to as "areas." These areas were divided into three lots, according as the male births in 1870 were less than 3000, between 3000 and 4000, or greater than 4000. When the districts in which alterations of boundaries occurring in the next five years were omitted, there were in the first group—Group A—45 districts, in the second—Group B—51, and in the third—Group C—17.

On account of the small number of deaths in the fifth year alone, the first three years and the next two (*i.e.* the fourth and fifth) were the periods taken. If the investigation were to be repeated the first two years and the next three would be the periods tried, but if selection is operative its existence is not likely to be completely hidden through such a small difference of arrangement. Accordingly the births, both male and female, in 1870 in each area were scheduled by adding those for the individual registration districts. The deaths of those under one year of age in 1870, between one and two in 1871, etc. were tabulated in the same way both for males and females. From these the deaths of those dying in the first three years of life, and also of those dying in the fourth and fifth could be ascertained. In addition, the total deaths under five years of age in each year (1870—1874) were taken out. Thus the "remaining deaths" (see § VIII) were found. The variables used were thus the same as those stated at the beginning of § IX, with 1870 read for 1903, etc. As before, it is claimed that by making x_0 and x_3 constant, we shall reach a fairly satisfactory "constant environment." To make it more satisfactory we can correct in addition for a constant population on which these "remaining deaths" in the five-year period are based. The labour involved in finding this population, however, is very great, and has been performed in the case of Group B only. Since in this case no marked alteration has been made in the correlation, it seems satisfactory for our present purpose to omit its consideration from the other groups.

This population can be obtained in the following manner. Those who died aged two in 1870 were taken from those who were born in 1869 minus those dying under one year in 1869 (our unit being the single year, as before). Those who died aged three in 1870 were taken from those born in 1868 minus those dying under one in 1868 minus those dying between one and two in 1869. And so on for four out of the five age-groups in each of the five years. In 1871, for example, we have to find the populations from which those who die between 0 and 1, 2 and 3, 3 and 4, and 4 and 5 are taken. The last of these will be those born in 1867 minus deaths (0—1) in 1867, minus deaths (1—2) in 1868, minus deaths (2—3) in 1869, minus deaths (3—4) in 1870. In all, then, the "remaining deaths" in the first five years of life will be based upon 20 distinct populations obtained in the above manner. If these are worked out and reduction performed it will be found that the complete population on which the whole of the "remaining deaths in the first five years" are based is:

Births 1866 + 2 (Births 1867) + 3 (Births 1868) + 4 (Births 1869)
+ Births 1874 + 2 (Births 1873) + 3 (Births 1872) + 4 (Births 1871)
− Deaths (age 0—1) {1866 + 2 (1867) + 3 (1868) + 4 (1869)
+ 1873 + 2 (1872) + 3 (1871)}
− Deaths (age 1—2) {1867 + 2 (1868) + 3 (1869) + 3 (1870) + 2 (1872) + 1873}
− Deaths (age 2—3) {1868 + 2 (1869) + 2 (1870) + 2 (1871) + 1873}
− Deaths (age 3—4) {1869 + 1870 + 1871 + 1872}.

These figures of births and deaths had to be obtained for each of the 51 areas forming Group B from those of the constituent registration districts from ten distinct annual reports. This extra variable will be denoted—in the case of Group B—by x_4.

(XIII) The statistical results reached from the various groupings are as follows:

Group A (45 *observations*).

Males			Females		
$\bar{x}_0 = 2701$	$\sigma_0 = 188\cdot 2$	$r_{01} = \cdot 275921$	$\bar{x}_0 = 2615$	$\sigma_0 = 198\cdot 6$	$r_{01} = \cdot 208638$
$\bar{x}_1 = 655$	$\sigma_1 = 120\cdot 0$	$r_{02} = \cdot 079396$	$\bar{x}_1 = 552$	$\sigma_1 = 103\cdot 8$	$r_{02} = \cdot 141500$
$\bar{x}_2 = 65$	$\sigma_2 = 24\cdot 2$	$r_{03} = \cdot 198703$	$\bar{x}_2 = 61$	$\sigma_2 = 22\cdot 4$	$r_{03} = \cdot 169066$
$\bar{x}_3 = 2922$	$\sigma_3 = 626\cdot 4$	$r_{12} = \cdot 509735$	$\bar{x}_3 = 2486$	$\sigma_3 = 568\cdot 6$	$r_{12} = \cdot 626603$
		$r_{13} = \cdot 887039$			$r_{13} = \cdot 874422$
${}_{3}r_{01} = +\cdot 220258$		$r_{23} = \cdot 681208$	${}_{3}r_{01} = +\cdot 127857$		$r_{23} = \cdot 737378$
${}_{3}r_{02} = -\cdot 077997$			${}_{3}r_{02} = +\cdot 025715$		
${}_{3}r_{12} = -\cdot 279682$,	${}_{03}r_{12} = -\cdot 269934 \pm \cdot 0932$		${}_{3}r_{12} = -\cdot 055465$,	${}_{03}r_{12} = -\cdot 059259 \pm \cdot 1002$	

Group B (51 *observations*).

Males			Females		
$\bar{x}_0 = 3409$	$\sigma_0 = 262{\cdot}7$	$r_{01} = {\cdot}435884$	$\bar{x}_0 = 3263$	$\sigma_0 = 281{\cdot}9$	$r_{01} = {\cdot}411332$
$\bar{x}_1 = 789$	$\sigma_1 = 148{\cdot}0$	$r_{02} = {\cdot}156272$	$\bar{x}_1 = 667$	$\sigma_1 = 145{\cdot}0$	$r_{02} = {\cdot}209618$
$\bar{x}_2 = 75$	$\sigma_2 = 28{\cdot}0$	$r_{03} = {\cdot}326415$	$\bar{x}_2 = 74$	$\sigma_2 = 29{\cdot}1$	$r_{03} = {\cdot}340572$
$\bar{x}_3 = 3484$	$\sigma_3 = 799{\cdot}2$	$r_{04} = {\cdot}915764$	$\bar{x}_3 = 2977$	$\sigma_3 = 718{\cdot}7$	$r_{04} = {\cdot}921920$
$\bar{x}_4 = 56853$	$\sigma_4 = 5077{\cdot}9$	$r_{12} = {\cdot}699475$	$\bar{x}_4 = 53772$	$\sigma_4 = 5039{\cdot}4$	$r_{12} = {\cdot}739053$
		$r_{13} = {\cdot}927496$			$r_{13} = {\cdot}928466$
		$r_{14} = {\cdot}162815$			$r_{14} = {\cdot}159766$
${}_3r_{12} = -{\cdot}205678$		$r_{23} = {\cdot}803504$	${}_3r_{12} = -{\cdot}290397$		$r_{23} = {\cdot}856040$
${}_{03}r_{12} = -{\cdot}148066 \pm {\cdot}0924$		$r_{24} = -{\cdot}007097$	${}_{03}r_{12} = -{\cdot}257776 \pm {\cdot}0882$		$r_{24} = {\cdot}062865$
${}_{034}r_{12} = -{\cdot}142090 \pm {\cdot}0925$		$r_{34} = {\cdot}101963$	${}_{034}r_{12} = -{\cdot}203430 \pm {\cdot}0905$		$r_{34} = {\cdot}122229$

Group C (17 *observations*).

Males			Females		
$\bar{x}_0 = 4396$	$\sigma_0 = 309{\cdot}9$	$r_{01} = {\cdot}689236$	$\bar{x}_0 = 4194$	$\sigma_0 = 350{\cdot}9$	$r_{01} = {\cdot}718320$
$\bar{x}_1 = 1246$	$\sigma_1 = 259{\cdot}3$	$r_{02} = {\cdot}432965$	$\bar{x}_1 = 1068$	$\sigma_1 = 257{\cdot}6$	$r_{02} = {\cdot}507559$
$\bar{x}_2 = 123$	$\sigma_2 = 41{\cdot}7$	$r_{03} = {\cdot}727943$	$\bar{x}_2 = 122$	$\sigma_2 = 43{\cdot}2$	$r_{03} = {\cdot}717773$
$\bar{x}_3 = 5846$	$\sigma_3 = 110{\cdot}2$	$r_{12} = {\cdot}759204$	$\bar{x}_3 = 4859$	$\sigma_3 = 101{\cdot}0$	$r_{12} = {\cdot}796579$
		$r_{13} = {\cdot}968693$			$r_{13} = {\cdot}942543$
${}_{03}r_{12} = -{\cdot}020180 \pm {\cdot}1636$		$r_{23} = {\cdot}782007$	${}_{03}r_{12} = +{\cdot}137224 \pm {\cdot}1605$		$r_{23} = {\cdot}826722$

Groups A, B and C together (113 *observations*).

Males			Females		
$\bar{x}_0 = 3276$	$\sigma_0 = 622{\cdot}6$	$r_{01} = {\cdot}808692$	$\bar{x}_0 = 3115$	$\sigma_0 = 596{\cdot}5$	$r_{01} = {\cdot}772263$
$\bar{x}_1 = 835$	$\sigma_1 = 250{\cdot}8$	$r_{02} = {\cdot}546280$	$\bar{x}_1 = 682$	$\sigma_1 = 230{\cdot}1$	$r_{02} = {\cdot}595457$
$\bar{x}_2 = 78$	$\sigma_2 = 35{\cdot}0$	$r_{03} = {\cdot}754827$	$\bar{x}_2 = 78$	$\sigma_2 = 35{\cdot}6$	$r_{03} = {\cdot}739053$
$\bar{x}_3 = 3561$	$\sigma_3 = 116{\cdot}0$	$r_{12} = {\cdot}793532$	$\bar{x}_3 = 3207$	$\sigma_3 = 107{\cdot}5$	$r_{12} = {\cdot}824615$
		$r_{13} = {\cdot}968835$			$r_{13} = {\cdot}953300$
${}_{03}r_{12} = -{\cdot}015226 \pm {\cdot}0634$		$r_{23} = {\cdot}837911$	${}_{03}r_{12} = +{\cdot}046446 \pm {\cdot}0634$		$r_{23} = {\cdot}864962$

(XIV) The following conclusions are to be drawn from the figures given in the previous paragraph.

(i) Although the correlation ${}_{03}r_{12}$ is significant in two cases only—for males in Group A and females in Group B—and then only just significant, yet this correlation is negative in all the four cases composing Groups A and B. Not much weight is to be attached to the figures given under Group C, since they are based on 17 observations only, and these mostly of the type which will be seen below to be of an unsatisfactory nature. The results for Groups A and B seem to point to a definite, though small, negative correlation between deaths in the first two years of life and deaths in the fourth and fifth, when allowance is made for environment.

(ii) The corresponding crude correlations for males and females are very similar throughout, and agree well within the probable error of either. The mean value of ${}_{03}r_{12}$ in Groups A and B for males is $-{\cdot}209$, and for females $-{\cdot}158$, but the higher value for males in Group A is accompanied by the lower value for females, while the reverse holds in Group B. At present then it is not possible to assert a significant difference between the correlations for males and females.

(iii) The very considerable labour involved in obtaining the figures necessary to use the extra variable in Group B has not been attended with any great alteration in the value of the correlation. It will be noticed that both for males and females x_4 is highly correlated with x_0 but only slightly with the other variables x_1, x_2 and x_3. Had x_4 been perfectly correlated with x_0, and perfectly uncorrelated with the other variables its introduction would have led to no alteration in the result. As it is, for males it has caused a reduction from $-{\cdot}148$ to $-{\cdot}142$, and for females from $-{\cdot}258$ to $-{\cdot}203$. Since ${}_{03}r_{12}$ is negative it can be seen from Yule's form of the partial correlation coefficient $\left(\text{viz. } {}_{034}r_{12} = \dfrac{{}_{03}r_{12} - {}_{03}r_{14} \cdot {}_{03}r_{24}}{\sqrt{1-{}_{03}r_{14}^2}\sqrt{1-{}_{03}r_{24}^2}}\right)$ that this reduction can only arise through ${}_{03}r_{14}$ and ${}_{03}r_{24}$ being of opposite sign. (The reverse proposition, that if ${}_{03}r_{14}$ and ${}_{03}r_{24}$ are of opposite sign there must be reduction, is not necessarily true.) *A priori* it would appear that if these correlations are of opposite sign, they will not be very large, from which it follows from the above expression that no great alteration in ${}_{03}r_{12}$ is to be expected by correcting for x_4. In the case of males, ${}_{03}r_{14}$ was $-{\cdot}5318$ and ${}_{03}r_{24}$ was $+{\cdot}0781$, while for females the corresponding values were $-{\cdot}4116$ and $+{\cdot}2373$. It is quite possible, of course, even if these correlations are of opposite sign, that the (negative) value of ${}_{034}r_{12}$ may be greater than the (negative) value of ${}_{03}r_{12}$. This will arise when the absolute value of ${}_{03}r_{14}$ is large and of ${}_{03}r_{24}$ small, or *vice versa*.

The omission of x_4, therefore, is not likely to vitiate any qualitative conclusion which may be drawn from a definitely significant value of ${}_{03}r_{12}$.

(iv) Had we employed the whole 113 observations in a single group we should have definitely concluded that selection was inoperative. The male and female results here are quite consistent with a zero correlation between deaths in the first three years of life and deaths in the fourth and fifth, when correction is made for environment. This is due to the fact that the whole country consists of a heterogeneous aggregation of more or less homogeneous subgroups. Relationships which exist for each of the individual subgroups may be quite hidden when these latter are clubbed together. We consider this resolution of the observations into homogeneous groupings to be very important, not only in the present problem, but in many others. The correlations found in the report previously referred to (§ IV) are probably affected in this way.

(XV) So far no attempt has been made to group the areas according to the conditions existing in them. Dividing them according to size may incidentally mean that one group has a preponderance of industrial or other kind of districts, but the main point of division has been in regard to the numbers of the births. In the present paragraph we shall attempt a division into rural and urban areas. It is not possible to place every registration district into an urban or rural category, and two selections were made. These were (1) the areas which would certainly under any process of allocation be termed rural, and (2) those which would similarly be termed urban. All the areas which contained considerable parts of both urban and rural districts, or which could not be easily placed in either category, were omitted. The nature of the division can be seen from Table II and the Registrar-General's Annual Report for 1870. We shall speak of (1) as being truly representative of rural, and (2) of urban districts. The data for each of these groups of areas can be taken out of Table II. The statistical constants were found to be:

(1) *Rural Districts* (30 *observations*).

1870 Cohort.

Males			Females		
$\bar{x}_0 = 3227$	$\sigma_0 = 488{\cdot}3$	$r_{01} = {\cdot}884159$	$\bar{x}_0 = 3090$	$\sigma_0 = 467{\cdot}8$	$r_{01} = {\cdot}839725$
$\bar{x}_1 = 644$	$\sigma_1 = 137{\cdot}0$	$r_{02} = {\cdot}466919$	$\bar{x}_1 = 531$	$\sigma_1 = 118{\cdot}9$	$r_{02} = {\cdot}537607$
$\bar{x}_2 = 51$	$\sigma_2 = 16{\cdot}0$	$r_{03} = {\cdot}853151$	$\bar{x}_2 = 48$	$\sigma_2 = 15{\cdot}8$	$r_{03} = {\cdot}817663$
$\bar{x}_3 = 2669$	$\sigma_3 = 581{\cdot}0$	$r_{12} = {\cdot}510967$	$\bar{x}_3 = 2240$	$\sigma_3 = 508{\cdot}0$	$r_{12} = {\cdot}571600$
		$r_{13} = {\cdot}929183$			$r_{13} = {\cdot}951313$
		$r_{23} = {\cdot}702179$			$r_{23} = {\cdot}708506$
${}_3r_{01} = +{\cdot}474158,$	${}_3r_{02} = -{\cdot}355781$		${}_3r_{01} = +{\cdot}220001,$	${}_3r_{02} = -{\cdot}102786$	
${}_3r_{12} = -{\cdot}537623,$	${}_{03}r_{12} = -{\cdot}448353 \pm {\cdot}0984$		${}_3r_{12} = -{\cdot}473133,$	${}_{03}r_{12} = -{\cdot}466612 \pm {\cdot}0963$	

(2) *Urban Districts* (55 *observations*).

1870 Cohort.

Males			Females		
$\bar{x}_0 = 3347$	$\sigma_0 = 713{\cdot}7$	$r_{01} = {\cdot}871381$	$\bar{x}_0 = 3222$	$\sigma_0 = 681{\cdot}7$	$r_{01} = {\cdot}857081$
$\bar{x}_1 = 929$	$\sigma_1 = 269{\cdot}5$	$r_{02} = {\cdot}573775$	$\bar{x}_1 = 802$	$\sigma_1 = 245{\cdot}5$	$r_{02} = {\cdot}657909$
$\bar{x}_2 = 99$	$\sigma_2 = 31{\cdot}6$	$r_{03} = {\cdot}867870$	$\bar{x}_2 = 97$	$\sigma_2 = 32{\cdot}8$	$r_{03} = {\cdot}881772$
$\bar{x}_3 = 4239$	$\sigma_3 = 1121{\cdot}1$	$r_{12} = {\cdot}730013$	$\bar{x}_3 = 3708$	$\sigma_3 = 1023{\cdot}0$	$r_{12} = {\cdot}809521$
		$r_{13} = {\cdot}963649$			$r_{13} = {\cdot}964803$
${}_3r_{12} = -{\cdot}073932$		$r_{23} = {\cdot}770623$	${}_3r_{12} = +{\cdot}044152$		$r_{23} = {\cdot}832386$
${}_{03}r_{12} = +{\cdot}005758 \pm {\cdot}0909$			${}_{03}r_{12} = +{\cdot}061815 \pm {\cdot}0909$		

(XVI) These figures at once show a marked difference between rural and urban districts. The male and female figures are consistent in each case. In rural districts there appears to be a very marked and significant negative correlation between deaths in the first three years of life and deaths in the fourth and fifth when correction is made for environment, while no relationship at all seems to hold for the urban areas. We can confidently say that selection was operative in the former, but are we to argue that it was entirely absent in the latter? Does selection prevail under rural conditions and not under urban? We think that no answer can be given to this question by means of the methods of this memoir. The practically zero correlations found for the urban areas are due, we believe, to external circumstances. There is considerable difference in the migratory habits of town and country populations. In industrial districts, and particularly those in the neighbourhood of London, migration is very marked. As an instance we can quote that at a Parliamentary bye-election in such a district of floating population in the middle of the present year it was stated that "on a register of 16,000 voters, over 3000 removals have to be traced, many of them outside the district." The migratory habits may be more strongly exercised now than in 1870; still, there was much movement of the population, especially in London, even then. We are fairly confident that in such a district a considerable proportion of the deaths of those of ages four and five in 1873 and 1874 respectively were not of the children born in that same district in 1870. The urban districts are not those in which the effects of selection can be detected by the methods we are using.

In the rural districts, however, the case is quite different. The areas dealt with comprise many registration districts and cover hundreds of square miles—in some cases whole counties. There is very little migration of married people in rural parts except at the michaelmas movings. Even this is on a small scale, and is rarely to a point outside the district, and can hardly ever be outside the area we are dealing with. We can affirm with boldness that for our rural areas the personnel of the child population aged five in 1874 was substantially the same as the infant population in 1870, and we believe the data for these rural areas to be the most satisfactory that can be obtained for the purpose we have in hand.

We have here very distinct evidence of selection, the correlations both for males and females approaching $-\cdot 5$. Even the correction brought about by including the variable x_4 (as in Group B, § XIII) is hardly likely to reduce this to a value which cannot be deemed significant. Even if ${}_{03}r_{14}$ and ${}_{03}r_{24}$ are of opposite signs and of magnitudes as high as $\cdot 5$, a value of ${}_{03}r_{12}$ equal to $-\cdot 5$ is reduced only to $-\cdot 33$ for ${}_{034}r_{12}$, still quite significant when the two groups of males and females are considered. It cannot from the figures we have used be said if there is or is not a selective death-rate in our larger towns. With the results for rural districts before us, and believing the data for the urban districts to be unsatisfactory, we see no reason for denying the operation of selection in industrial districts. The correlation observed in Group C (§ XIII above) may be accounted for by the fact that almost all the districts used were urban, while in Groups A and B (particularly the latter) the rural ones occurred.

We conclude then that in any further investigation into the prevalence of selection in man by the method of this memoir rural communities only should be chosen for study.

(XVII) As the two samples (not really independent) of males and females used above contained only 30 observations each they are hardly sufficient to enable definite assertions of the intensity of selection to be made, and accordingly further samples were taken. These were obtained by considering those born, both males and females, in the years 1871 and 1872 in the 30 rural districts, and following them through the first five years of existence as before. Using a similar notation to that in § IX, the constants deduced for the 1871 cohort were:

Rural Districts (30 *observations*).

1871 Cohort.

Males			Females		
$\bar{x}_0 = 3226$	$\sigma_0 = 506{\cdot}6$	$r_{01} = {\cdot}853800$	$\bar{x}_0 = 3114$	$\sigma_0 = 480{\cdot}8$	$r_{01} = {\cdot}840147$
$\bar{x}_1 = 602$	$\sigma_1 = 137{\cdot}6$	$r_{02} = {\cdot}522615$	$\bar{x}_1 = 491$	$\sigma_1 = 108{\cdot}9$	$r_{02} = {\cdot}465796$
$\bar{x}_2 = 60$	$\sigma_2 = 13{\cdot}8$	$r_{03} = {\cdot}835691$	$\bar{x}_2 = 59$	$\sigma_2 = 18{\cdot}2$	$r_{03} = {\cdot}797256$
$\bar{x}_3 = 2635$	$\sigma_3 = 582{\cdot}3$	$r_{12} = {\cdot}576331$	$\bar{x}_3 = 2205$	$\sigma_3 = 527{\cdot}2$	$r_{12} = {\cdot}597272$
		$r_{13} = {\cdot}944432$			$r_{13} = {\cdot}956679$
		$r_{23} = {\cdot}706543$			$r_{23} = {\cdot}699196$
$_3r_{01} = +{\cdot}357550$,	$_3r_{02} = -{\cdot}174544$		$_3r_{01} = +{\cdot}440569$,	$_3r_{02} = -{\cdot}212350$	
$_3r_{12} = -{\cdot}390993$,	$_{03}r_{12} = -{\cdot}357369 \pm {\cdot}1074$		$_3r_{12} = -{\cdot}344148$,	$_{03}r_{12} = -{\cdot}285684 \pm {\cdot}1131$	

A comparison of these figures with those given for rural districts in § XV, shows that the crude correlations have not significantly altered in passing from one year to the next. The intensity of selection, as exhibited by the correlation $_{03}r_{12}$, however, is smaller for the latter year. Paying regard to the probable errors, we see that the results for males in the two years are quite consistent with an actual value of $_{03}r_{12}$ midway between those found, and similarly for the females. Attention must be drawn, too, to another important fact. The environmental conditions have not greatly changed from the one cohort to the other, but the mean value of the number of deaths in the earlier one is considerably larger than the number for the later. This is followed, however, by the deaths in the second period being greater for the second cohort than the first. This is in itself distinct evidence of selection. If a greater proportion of the cohort is killed off in the first period, a smaller proportion is killed off in the second, and this is more than can be accounted for by the fact that a smaller population remains to be aimed at in the second period. (See § XVIII.) This may be the cause of the alteration in the magnitude of $_{03}r_{12}$. For when more are killed off in the first period, selection is probably more stringent, and larger correlations may, therefore, be expected.

Some peculiar differences in the male and female variabilities may be noticed. In going from the first cohort to the second the female variability for deaths in the first period has fallen considerably, but has risen in the second period. For males, however, it has risen (slightly) in the first period and fallen in the second. The female environment, too, is more variable for the second cohort than the first, but that for the males has remained practically unchanged.

(XVIII) The corresponding results for the 1872 cohort were:

Rural Districts (30 *observations*).

1872 Cohort.

Males			Females		
$\bar{x}_0 = 3291$	$\sigma_0 = 520\cdot3$	$r_{01} = \cdot797697$	$\bar{x}_0 = 3150$	$\sigma_0 = 477\cdot0$	$r_{01} = \cdot814548$
$\bar{x}_1 = 588$	$\sigma_1 = 136\cdot8$	$r_{02} = \cdot608166$	$\bar{x}_1 = 494$	$\sigma_1 = 120\cdot2$	$r_{02} = \cdot595129$
$\bar{x}_2 = 62$	$\sigma_2 = 21\cdot9$	$r_{03} = \cdot803398$	$\bar{x}_2 = 61$	$\sigma_2 = 18\cdot5$	$r_{03} = \cdot753803$
$\bar{x}_3 = 2634$	$\sigma_3 = 605\cdot9$	$r_{12} = \cdot745228$	$\bar{x}_3 = 2189$	$\sigma_3 = 545\cdot1$	$r_{12} = \cdot711154$
		$r_{13} = \cdot951885$			$r_{13} = \cdot971880$
		$r_{23} = \cdot828195$			$r_{23} = \cdot795986$
${}_3r_{01} = +\cdot180609$,	${}_3r_{02} = -\cdot171419$		${}_3r_{01} = +\cdot529577$,	${}_3r_{02} = -\cdot012287$	
${}_3r_{12} = -\cdot251059$,	${}_{03}r_{12} = -\cdot227143 \pm \cdot1168$		${}_3r_{12} = -\cdot438124$,	${}_{03}r_{12} = -\cdot508864 \pm \cdot0912$	

Although the absolute value of ${}_{03}r_{12}$ here is more than twice as great for females as for males, the difference is not so great as to render certain a significant difference between the sexes. The negative value which is here reached for females is the largest yet obtained. This only arises through the high positive value of ${}_3r_{01}$ and very small negative value of ${}_3r_{02}$, the (negative) value of ${}_3r_{12}$ not being so large as for the 1870 cohort.

The mean of the three correlations for males is $-\cdot3433$ and for females $-\cdot4204$, and this difference implies no distinction in the intensity of selection, especially as the correlations for females are not uniformly higher than those for males. We need not necessarily expect that the correlations for the two sexes found by our present methods should be equal. It is quite conceivable that the best periods of measurement for boys are not the same as for girls. The sex incidence of death in the early years of life is well known, and as many boys die in the first three years of life as girls in the first six years. The slower rate at which the girls are killed off renders it advisable to have a longer second period for girls than for boys. It is interesting to note that the three values of the correlations for males progress in the same way as the mean numbers of those killed off in the first period. Were we arguing from the male figures alone we could make a strong point of this fact that when the deaths in the first three years of life are greatest there is the largest (negative) relationship between the numbers of these deaths and those in the next two years of life. For it most certainly indicates that many of those deaths in the first three years must be due to selective action. The female figures are not quite so consistent, and this may be, as stated above, because the periods used are not so well adapted for the females as for the males. Still, the same kind of relationship as for the males can be detected for the cohorts of females born in the separate pairs of years 1870 and 1871, and 1871 and 1872, though not for 1870 and 1872.

Another point to be noticed from the three sets of figures is the considerable variation in the values of σ_2, particularly for the males. The coefficient of variation of x_2 is, for males, ·33, ·23 and ·35 for the successive years, the corresponding figures for the females being ·33, ·31 and ·30. These point to considerable local differences in the incidence of death in the fourth and fifth years of life, which vary from year to year, and can be applied as an additional argument against the use of the whole country in successive years as a means to detect selection (see § IV). The large values of σ_2 for the 1872 cohort were due to large increases in the magnitudes of that variable over those of the previous years in some areas in the north of England, those for most of the remaining areas showing small decreases.

Some instructive information can be obtained from the mean values of the deaths given for the three cohorts by reducing these to rates, those for the first period being expressed on the number born, and those for the second on the number of children surviving the first period. We find the following figures:

	1870 Cohort		1871 Cohort		1872 Cohort	
	Males	Females	Males	Females	Males	Females
First Period	·1996	·1718	·1866	·1541	·1787	·1568
Second Period ...	·01974	·01876	·02287	·02249	·02294	·02297

For males, the highest rate in the first period is followed by the lowest in the second, and conversely. For females, the former statement also holds good, but the converse is not definitely brought out. The difference between the female figures of the 1871 and 1872 cohorts is, however, very small.

(XIX) Larger correlations might be expected if we could go beyond the first five years of life. In the English rural districts there seems no valid reason why the present method could not be applied up to the age of ten, or even fifteen. There is little migration from rural areas up to the latter age, and the deaths in each year of life up to fifteen would be exceedingly valuable. Unfortunately, however, such information is not given in any of the reports of the Registrar-General, not even the earlier ones, and so there is no means of dealing with England and Wales. In Sweden, deaths at each age of life have long been published, but only for the aggregate of rural districts and of urban districts. As stated elsewhere, it is not satisfactory to deal with such aggregations. But in the annual volumes of the *Prussian Statistics* there is an admirable storehouse of vital statistics ready for our purpose. The deaths—both male and female—are given each year at every year of age for the rural as well as the urban portions of each province. For the purpose of a scientific enquiry like the present no collection of statistics of whose existence we are aware approaches in value the Prussian vital statistics.

There are in all 36 provinces. One, however,—Berlin—has no rural portion and another—Sigmaringen—is too small to use. Of the remainder, four were omitted which showed signs (as exhibited by rapidly increasing numbers of births) of considerable immigration, and the rural portions of the remaining 30 taken for the investigation. The year 1881 was taken as the starting point of the work, and enough data were extracted to follow those born in that year, both male and female, throughout the first twenty years of their life. The mode of procedure was similar to that adopted for the English rural areas, and the same kind of measure of environment was employed, viz., in addition to making the births in the year 1881 constant, to fix also the "remaining deaths" (see § VIII) in the interval composing the two separate periods considered. It is possible that when this interval is as long as ten years, the inclusion of the additional variable x_4—the population on which these "remaining deaths" are based—will make greater alteration than in the case when its effect was actually investigated (§ XIII). The labour of working this extra variable out in such a case, however, is almost prohibitive, and we shall content ourselves with ascertaining its probable effect in the case of those correlations whose value we emphasise.

XX. The first work attempted on the Prussian data was to investigate the effects on the second five years of life of variations in the numbers of deaths in the first five. The variables used were accordingly:

x_0 = Births in the year 1881,
x_1 = Deaths in the first five years of life of those born in 1881,
x_2 = ,, ,, ,, second ,, ,, ,, ,, ,, ,, ,, ,, ,,
x_3 = Remaining deaths in the ten years 1881–90.

As the number of observations was again small, a similar study was made of the figures for those born in 1882. The complete data for both years will be found collected in Tables V and VI. The means, standard deviations and correlations for each cohort, both for male and female births, are collected below.

Prussian Rural Districts (30 *observations*).

1881 Cohort.

Males			Females		
$\bar{x}_0 = 9407$	$\sigma_0 = 4586 \cdot 9$	$r_{01} = \cdot 971501$	$\bar{x}_0 = 8917$	$\sigma_0 = 4398 \cdot 1$	$r_{01} = \cdot 964662$
$\bar{x}_1 = 2760$	$\sigma_1 = 1794 \cdot 5$	$r_{02} = \cdot 872522$	$\bar{x}_1 = 2387$	$\sigma_1 = 1558 \cdot 6$	$r_{02} = \cdot 892746$
$\bar{x}_2 = 239$	$\sigma_2 = 125 \cdot 5$	$r_{03} = \cdot 971583$	$\bar{x}_2 = 239$	$\sigma_2 = 127 \cdot 8$	$r_{03} = \cdot 968570$
$\bar{x}_3 = 28277$	$\sigma_3 = 18056 \cdot 8$	$r_{12} = \cdot 842849$	$\bar{x}_3 = 24650$	$\sigma_3 = 15723 \cdot 8$	$r_{12} = \cdot 873589$
		$r_{13} = \cdot 998637$			$r_{13} = \cdot 998518$
		$r_{23} = \cdot 853132$			$r_{23} = \cdot 878454$
${}_{3}r_{01} = + \cdot 100518$,	${}_{3}r_{02} = + \cdot 353349$		${}_{3}r_{01} = - \cdot 182603$,	${}_{3}r_{02} = + \cdot 352540$	
${}_{3}r_{12} = - \cdot 334884$,	${}_{03}r_{12} = - \cdot 398024 \pm \cdot 1036$		${}_{3}r_{12} = - \cdot 137095$,	${}_{03}r_{12} = - \cdot 079038 \pm \cdot 1224$	

1882 Cohort.

Males			Females		
$\bar{x}_0 = 9297$	$\sigma_0 = 4689\cdot5$	$r_{01} = \cdot971294$	$\bar{x}_0 = 8793$	$\sigma_0 = 4504\cdot8$	$r_{01} = \cdot967380$
$\bar{x}_1 = 2938$	$\sigma_1 = 1981\cdot5$	$r_{02} = \cdot918784$	$\bar{x}_1 = 2511$	$\sigma_1 = 1716\cdot2$	$r_{02} = \cdot929847$
$\bar{x}_2 = 217$	$\sigma_2 = 107\cdot9$	$r_{03} = \cdot975529$	$\bar{x}_2 = 223$	$\sigma_2 = 103\cdot3$	$r_{03} = \cdot978221$
$\bar{x}_3 = 28137$	$\sigma_3 = 18015\cdot8$	$r_{12} = \cdot858415$	$\bar{x}_3 = 24480$	$\sigma_3 = 15645\cdot2$	$r_{12} = \cdot861304$
		$r_{13} = \cdot996780$			$r_{13} = \cdot996398$
		$r_{23} = \cdot861457$			$r_{23} = \cdot876992$
${}_3r_{01} = -\cdot062050$,	${}_3r_{02} = +\cdot702213$		${}_3r_{01} = -\cdot415700$,	${}_3r_{02} = +\cdot721448$	
${}_3r_{12} = -\cdot006614$,	${}_{03}r_{12} = +\cdot052016 \pm \cdot1228$		${}_3r_{12} = -\cdot307485$,	${}_{03}r_{12} = -\cdot012034 \pm \cdot1231$	

The above correlations are not very regular, and do not, in themselves, give very definite evidence of selection. Had the 1881 figures alone been used we might have been tempted to state that they indicated some small selective action. Three out of the four, however, point to no relationship at all between the deaths in the first five years of life, and those in the next five. There is no reason why we should get the same results for the two different cohorts. Selection is probably not uniform in its action on the separate child populations, and the use of those periods which give significant correlations in the one case may be inappropriate to the other. The conditions may be such that diphtheria, fever and measles strike at the first cohort throughout its first five years, but at the other only throughout its first four. Thus if two five-years periods are best for the one, one of four and the other of six may be best for the other. By taking five and five for the latter we get an overlapping which will cause a diminution in the correlation.

A study of the mean values of the variable for the two distinct cohorts, however, gives us evidence of selection. Although the mean number of births diminishes from the first to the second, the mean number of deaths in the first five years of life increases, and this is followed by a decrease in the mean number of deaths in the second five years of life, the environmental conditions remaining practically the same. Expressed in the form of rates, based upon the number born for the first period and the number surviving to the age of five for the second, we find the following figures:

	1881 Cohort		1882 Cohort	
	Males	Females	Males	Females
First Period	·2934	·2677	·3160	·2856
Second Period ...	·0359	·0366	·0341	·0355

Thus, both for males and females, those belonging to the cohort in which the mortality in the first period is high, and who survive that period have a greater chance of surviving the second period also. In other words, the survivors of the more stringent early period are better fitted for their after life than the survivors of the less stringent.

(XXI) The division into five-year periods in the last paragraph was a purely arbitrary one and has no statistical or other reason to justify it. In fact, the results showed that the division was a bad one and tended to hide the effects of selection by the overlapping of the natural periods in which selection may be expected to show itself. These natural periods were brought back to mind by the re-reading of Pearson's *The Chances of Death.* He has there (Vol. I. p. 34) shown that there is a mortality of childhood quite distinct from that of infancy, almost all the latter occurring in the first two years of life. In the case of the two chief foes of child-life—scarlet fever and diphtheria—the attack is not so much in the first two years of life as in the third and later years. Pearson's figures showed that the mean age of death (for males) in the childhood period was 6, with a standard deviation of 3·5. Accordingly we chose the first two years of life to be the one period, and the next eight to be the second. In this way we are investigating the effects on the mortality of "childhood" of variations in the mortality of infancy. The data used in this case can be quickly deduced from the third, fourth and fifth columns of figures for males and females in Tables V and VI. The notation used was:

x_0 = Births,
x_1 = Deaths in first two years of life,
x_2 = ,, ,, next eight ,, ,, ,,
x_3 = Remaining deaths in the ten years.

As before the work was performed on the figures for those born in 1882 as well as for those born in 1881, and the results are:

1881 *Cohort* (30 *observations*).

Males			Females		
$\bar{x}_0 = 9407$	$\sigma_0 = 4586{\cdot}9$	$r_{01} = {\cdot}966797$	$\bar{x}_0 = 8917$	$\sigma_0 = 4398{\cdot}1$	$r_{01} = {\cdot}963495$
$\bar{x}_1 = 2270$	$\sigma_1 = 1502{\cdot}2$	$r_{02} = {\cdot}867956$	$\bar{x}_1 = 1914$	$\sigma_1 = 1276{\cdot}4$	$r_{02} = {\cdot}871859$
$\bar{x}_2 = 729$	$\sigma_2 = 461{\cdot}0$	$r_{03} = {\cdot}971578$	$\bar{x}_2 = 711$	$\sigma_2 = 445{\cdot}0$	$r_{03} = {\cdot}968570$
$\bar{x}_3 = 28277$	$\sigma_3 = 18056{\cdot}8$	$r_{12} = {\cdot}825960$	$\bar{x}_3 = 24650$	$\sigma_3 = 15723{\cdot}8$	$r_{12} = {\cdot}854617$
		$r_{13} = {\cdot}991230$			$r_{13} = {\cdot}991070$
		$r_{23} = {\cdot}889151$			$r_{23} = {\cdot}909771$
${}_3r_{01} = +{\cdot}119558$,	${}_3r_{02} = +{\cdot}037627$		${}_3r_{01} = +{\cdot}107783$,	${}_3r_{02} = -{\cdot}090243$	
${}_3r_{12} = -{\cdot}916019$,	${}_{03}r_{12} = -{\cdot}927836 \pm {\cdot}0171$		${}_3r_{12} = -{\cdot}849620$,	${}_{03}r_{12} = -{\cdot}848276 \pm {\cdot}0345$	

1882 *Cohort* (30 *observations*).

Males			Females		
$\bar{x}_0 = 9297$	$\sigma_0 = 4689\cdot5$	$r_{01} = \cdot967684$	$\bar{x}_0 = 8793$	$\sigma_0 = 4504\cdot8$	$r_{01} = \cdot965285$
$\bar{x}_1 = 2424$	$\sigma_1 = 1654\cdot1$	$r_{02} = \cdot895378$	$\bar{x}_1 = 2004$	$\sigma_1 = 1397\cdot1$	$r_{02} = \cdot913858$
$\bar{x}_2 = 732$	$\sigma_2 = 470\cdot7$	$r_{03} = \cdot975529$	$\bar{x}_2 = 729$	$\sigma_2 = 449\cdot9$	$r_{03} = \cdot978221$
$\bar{x}_3 = 28137$	$\sigma_3 = 18015\cdot8$	$r_{12} = \cdot860826$	$\bar{x}_3 = 24480$	$\sigma_3 = 15645\cdot2$	$r_{12} = \cdot883242$
		$r_{13} = \cdot995812$			$r_{13} = \cdot995626$
		$r_{23} = \cdot890463$			$r_{23} = \cdot911441$
${}_{3}r_{01} = -\cdot186994$,	${}_{3}r_{02} = +\cdot266908$		${}_{3}r_{01} = -\cdot221147$,	${}_{3}r_{02} = +\cdot260740$	
${}_{3}r_{12} = -\cdot622720$,	${}_{03}r_{12} = -\cdot605038 \pm \cdot0781$		${}_{3}r_{12} = -\cdot629876$,	${}_{03}r_{12} = -\cdot607764 \pm \cdot0770$	

The results obtained here are remarkable. There is a significant difference in the values of ${}_{03}r_{12}$ for the two cohorts, but between males and females for each the difference is not significant. The (negative) correlations are quite the highest we have obtained and may be put roughly at $-\cdot9$ for those born in 1881 and $-\cdot6$ for those born in 1882. It is hardly likely that these figures would be much reduced by further correcting for the actual populations on which the deaths x_3 are based. The expression

$${}_{034}r_{12} = \frac{{}_{03}r_{12} - {}_{03}r_{14} \cdot {}_{03}r_{24}}{\sqrt{1 - {}_{03}r_{14}^{\,2}} \cdot \sqrt{1 - {}_{03}r_{24}^{\,2}}}$$

at once shows that considerable reduction can only arise by having ${}_{03}r_{14}$ and ${}_{03}r_{24}$ of opposite signs and of absolute magnitudes over ·5 each. If ${}_{03}r_{14} = \pm\cdot5$ and ${}_{03}r_{24} = \mp\cdot5$, values of ${}_{03}r_{12}$ of $-\cdot9$ and $-\cdot6$ give rise to values of ${}_{034}r_{12}$ equal to $-\cdot867$ and $-\cdot467$. For higher absolute values of ${}_{03}r_{14}$ and ${}_{03}r_{24}$ the reduction largely increases (${}_{034}r_{12}$ may even become positive), but it seems very improbable that such high values of opposite sign can exist (see also § XIV). The results found in the present paragraph then must indicate considerable selection in the infant period. For the 1881 cohort, indeed, the distribution of the childhood mortality was almost completely settled by the knowledge of the distribution of the infantile mortality. Those born in 1882 seemed to have been constitutionally weaker than those born in the earlier year, for the total deaths of the former in the first ten years of life was greater by 7950, though the number born was 7020 less and the environmental conditions practically unchanged.

The reason of the small correlations in the previous paragraph is now evident. The interval embraced in the first five years of life overlaps the two periods of infancy and childhood. The effects of infantile selection are already apparent by the third, fourth and fifth years, so that in districts of high (corrected) mortality in the first two years there is low (corrected) mortality in the three mentioned, and *vice versa*. This will exert a cancelling effect on the deviations from the respective means which are of opposite sign. The results of this paragraph confirm in a striking manner the conclusions of Pearson, derived from a study of the theoretical frequency distribution of mortality, on the distinction between infantile and child mortality.

A study of the means does not in this case give any definite evidence of selection. Expressed as rates, as before, we find

	1881 Cohort		1882 Cohort	
	Males	Females	Males	Females
First Period	·2413	·2146	·2607	·2279
Second Period ...	·1021	·1013	·1079	·1074

Thus the higher rate in the first period is followed by a higher rate in the second in each case.

(XXII) One further case was attempted, in order to discover, if possible, the effects of selection at higher ages. In order to do this, the deaths in the first five years of life were correlated with those in the next fifteen, the other variable, in addition to the births, being the "remaining deaths" (see § XII) in the twenty years considered. We cannot lay much stress on the results found here for two reasons, (*a*) the effects of migration are hardly likely to be negligible, even in rural districts, up to the twentieth year of life, and (*b*) the further correction necessary for the constancy of the population on which the "remaining deaths" are based is far more important than when only a few of the early years of life are considered. The data are reduced and collected in Tables V and VI, and the notation used was:

$x_0 =$ Births,

$x_1 =$ Deaths in the first five years of life,

$x_2 =$ Deaths in the next fifteen years of life,

$x_3 =$ "Remaining Deaths" in the twenty years.

The following results were found:

1881 *Cohort* (30 *observations*).

Males			Females		
$\bar{x}_0 = 9407$	$\sigma_0 = 4586{\cdot}9$	$r_{01} = \cdot 971501$	$\bar{x}_0 = 8917$	$\sigma_0 = 4398{\cdot}1$	$r_{01} = \cdot 964164$
$\bar{x}_1 = 2760$	$\sigma_1 = 1794{\cdot}5$	$r_{02} = \cdot 922620$	$\bar{x}_1 = 2387$	$\sigma_1 = 1558{\cdot}6$	$r_{02} = \cdot 941523$
$\bar{x}_2 = 407$	$\sigma_2 = 180{\cdot}3$	$r_{03} = \cdot 975243$	$\bar{x}_2 = 420$	$\sigma_2 = 188{\cdot}3$	$r_{03} = \cdot 974481$
$\bar{x}_3 = 60940$	$\sigma_3 = 37130{\cdot}2$	$r_{12} = \cdot 869830$	$\bar{x}_3 = 53303$	$\sigma_3 = 32245{\cdot}3$	$r_{12} = \cdot 885565$
		$r_{13} = \cdot 995397$			$r_{13} = \cdot 994034$
		$r_{23} = \cdot 870802$			$r_{23} = \cdot 893501$
${}_3r_{01} = + \cdot 035249,$	${}_3r_{02} = + \cdot 326270$		${}_3r_{01} = - \cdot 183920,$	${}_3r_{02} = + \cdot 702601$	
${}_3r_{12} = + \cdot 064438,$	${}_{03}r_{12} = + \cdot 059842 \pm \cdot 1227$		${}_3r_{12} = - \cdot 053205,$	${}_{03}r_{12} = + \cdot 155386 \pm \cdot 1202$	

1882 *Cohort* (30 *observations*).

Males			Females		
$\bar{x}_0 = 9297$	$\sigma_0 = 4689{\cdot}5$	$r_{01} = {\cdot}971294$	$\bar{x}_0 = 8793$	$\sigma_0 = 4504{\cdot}8$	$r_{01} = {\cdot}967380$
$\bar{x}_1 = 2938$	$\sigma_1 = 1981{\cdot}5$	$r_{02} = {\cdot}933874$	$\bar{x}_1 = 2511$	$\sigma_1 = 1716{\cdot}2$	$r_{02} = {\cdot}919096$
$\bar{x}_2 = 400$	$\sigma_2 = 170{\cdot}0$	$r_{03} = {\cdot}975913$	$\bar{x}_2 = 413$	$\sigma_2 = 167{\cdot}7$	$r_{03} = {\cdot}980276$
$\bar{x}_3 = 60750$	$\sigma_3 = 37057{\cdot}4$	$r_{12} = {\cdot}856059$	$\bar{x}_3 = 53053$	$\sigma_3 = 32144{\cdot}1$	$r_{12} = {\cdot}829052$
		$r_{13} = {\cdot}990449$			$r_{13} = {\cdot}983029$
		$r_{23} = {\cdot}870171$			$r_{23} = {\cdot}874315$
$_3r_{01} = +{\cdot}156320,$	$_3r_{02} = +{\cdot}786666$		$_3r_{01} = +{\cdot}103156,$	$_3r_{02} = +{\cdot}642622$	
$_3r_{12} = -{\cdot}085386,$	$_{03}r_{12} = -{\cdot}341689 \pm {\cdot}1088$		$_3r_{12} = -{\cdot}341704,$	$_{03}r_{12} = -{\cdot}552700 \pm {\cdot}0855$	

Although the male and female figures are consistent within each cohort, those for the two cohorts are quite incompatible. For the earlier one no evidence of selection appears; for the latter, however, significant negative correlations are found. It is difficult to suggest a reason for the negative correlations in this case apart from selective action. If the argument be advanced that migration (which we believe must be quite appreciable when dealing with such a long period) may account for it, it must then be shown that migration is differential in its operation, *i.e.* that in districts of high mortality in the first five years of life *emigration* occurs while in the districts of low mortality in those years *immigration* takes place. In the absence of any evidence in favour of this differentiation we must conclude that selection was operative on the 1882 cohort up to the twentieth year.

This apparent inconsistency between the results for the two cohorts may be explained if we suppose that selection is irregular in its action. We found a higher intensity of selection in the first ten years of life for the 1881 cohort than for the 1882. It may be, therefore, that selection exhausted itself for the former in the first ten years, and thus we obtained small positive values for $_{03}r_{12}$ when proceeding to the twenty years. On the other hand the smaller values for the first ten years of life for the second cohort are followed by quite significant values for the twenty years. Thus selection in this case was reserved for the later years. This may not be the true explanation of the differences to be noted in the figures in this paragraph, but it seems a very probable one. Those born in Prussian rural districts in 1881 felt almost the full effects of selection in the first ten years of their life, but many of those born in 1882 felt its operation in later years. Much light would be thrown on the subject by taking different periods—in particular, the correlation between the deaths in the first two years of life with those in the next eighteen would be instructive—but the work has already far exceeded the original limits set for it and further investigation must be postponed.

Expressing the mean values of x_1 and x_2 in terms of rates (see § XVIII) we find the following figures:

	First Cohort		Second Cohort	
	Males	Females	Males	Females
First Period	·2934	·2677	·3160	·2856
Second Period ...	·0612	·0643	·0629	·0657

As in the last case, these figures, in themselves, give no evidence of the action of selection.

XXIII. No other correlations have been worked out. Sufficient has been done to show conclusively that there is a definitely significant negative correlation between the number of deaths in the early years of life and those in a subsequent period, when correction is made for environment. It would be easy to suggest all sorts of different arrangements of the data and of the periods of time in which selection is taken to act. In particular it would be instructive to ascertain the correlations for many different periods when working with the same data. It is quite probable that we should have got higher values in our work on the English rural districts by taking the first two years of life with the next three, instead of the first three with the next two. But this as well as many other improvements and extensions must be left for future investigation.

Further light, however, is thrown on the subject with little extra labour by means of the various partial regression coefficients. Since it is useful to know also the partial standard deviations, these partial regressions have been worked out by Yule's method (and verified by the determinant method). With an obvious notation, we know that the partial regression ${}_{03}b_{21} = {}_{03}r_{12} \cdot \frac{{}_{03}\sigma_2}{{}_{03}\sigma_1}$. These standard deviations and regressions have been worked out in these cases for which the data is the most satisfactory, viz. English and Prussian rural districts. The regressions are given in the last column of the table below.

When correction has been made for environment the distribution of the variable x_1 among the districts may be considered to be approximately Gaussian. Taking the case of the males in 1870, it follows that when the environmental conditions are rendered uniform throughout the areas, the range of the variable x_1 will be from 510 to 778 approximately, *this difference arising solely from random causes.* In those districts for which the value of x_1 was 644, the most likely value of x_2 is 51, and the partial regressions tell us that for those districts in which the value of x_1 was 544 the most likely value of x_2 was 62, while if x_1 had the value 744 the most likely value of x_2 was 40. This shows in an emphatic manner that when allowance has been made for environment, the districts in which the deaths in the first period were above the average were, on the whole, those districts in which the deaths in the second period were below the average, and conversely. For the females of 1870, we find that districts in which the (corrected) deaths in the first period were 431, 531 and 631 had on the average, 63, 48 and 33 deaths, respectively, in the second period. For the Prussian males of 1881, values of x_1 equal to 2170, 2270 and 2370 were followed, on the average, by values of x_2 equal to 827, 729 and 631 when environmental conditions were allowed for. For the females, deaths in the first period of numbers 1814, 1914 and 2014 were followed, on the whole, by 803, 711 and 619 deaths respectively in the second period.

Partial Regressions and other Constants for the Data from English and Prussian Rural Districts.

Data		$\bar{x}_0$	$\bar{x}_1$	$\bar{x}_2$	${}_{03}\sigma_1$	${}_{03}\sigma_2$	$\frac{{}_{03}\sigma_2}{{}_{03}\sigma_1}$	${}_{03}r_{12}$	Expected correlation if no selection	Partial regression x_2 on x_1
Males: English Rural Districts,	1870	3227	644	51	44·582	10·634	·2385	−·4483	−·0828	−·1070
	1871	3226	602	60	42·247	9·589	·2270	−·3574	−·1014	−·0811
	1872	3291	588	62	41·220	12·108	·2938	−·2271	−·0807	−·0667
Prussian Rural Districts	1881	9407	2270	729	198·042	210·803	1·0644	−·9278	−·0958	−·9875
	1882	9297	2424	732	148·377	206·413	1·3911	−·6050	−·0765	−·8417
Females: English Rural Districts,	1870	3090	531	48	34·340	11·126	·3240	−·4666	−·0708	−·1512
	1871	3114	491	59	28·468	12·736	·4474	−·2857	−·0505	−·1278
	1872	3150	494	61	24·011	11·175	·4654	−·5089	−·0496	−·2368
Prussian Rural Districts	1881	8917	1914	711	169·182	183·986	1·0875	−·8483	−·0933	−·9225
	1882	8793	2004	729	117·355	178·700	1·5227	−·6078	−·0705	−·9255

It will be noticed that, except in the case of the Prussian data for 1881, the partial regressions are always greater for the females than for the males, this arising chiefly from the considerably smaller partial standard deviations of the deaths in the first period in the case of the females. At the same time the mean values of the number of deaths in the first period are much smaller for the females. Thus the effect on the number of deaths in the second period of an excess or defect from the mean number in the first period is greater for females than for males even though the mean number for the females is smaller. This points to a greater intensity of selection for females than for males, when measured on the periods we have been considering. Thus the consequence of an epidemic killing off both boys and girls from a population under the age of three, is that the girls are left more immune in the fourth and fifth years of life than are the boys. The effect at later ages may, of course, be reversed, but our data for English rural districts point definitely to this conclusion for the first five years.

(XXIV) We can consider now the point raised before (§ V), that when correction has been made for environment and size of cohort, some negative correlation between the deaths in the periods used might be expected even if selection were inoperative; this negative correlation arising from the fact that the available population at which death may strike in the second period will be larger or smaller according as the deaths in the first period were few or many. The magnitude of this negative correlation can be found as follows: if B is the number of births, d_1 and d_2 the deaths in the first and second periods respectively, the death-rate in the second period is $\frac{d_2}{B-d_1}$ and this equals $\frac{\bar{x}_2}{\bar{x}_0-\bar{x}_1}$. In a district for which the number of deaths in the first period exceeded the mean by the standard deviation (${}_{03}\sigma_1$), the available population for the second period would be $\bar{x}_0-\bar{x}_1-{}_{03}\sigma_1$, and the most probable number of deaths in this district if there is no selection, therefore, would be $\frac{\bar{x}_2}{\bar{x}_0-\bar{x}_1}(\bar{x}_0-\bar{x}_1-{}_{03}\sigma_1)=\bar{x}_2-\frac{{}_{03}\sigma_1\bar{x}_2}{\bar{x}_0-\bar{x}_1}$. Thus a positive deviation of ${}_{03}\sigma_1$ in the first period would be followed, *on the average*, by a negative one of $\frac{{}_{03}\sigma_1\bar{x}_2}{\bar{x}_0-\bar{x}_1}$ in the second. The mean value of the (negative) product moment is thus $\frac{\bar{x}_2\cdot{}_{03}\sigma_1^2}{\bar{x}_0-\bar{x}_1}$, and the correlation $-\frac{\bar{x}_2}{\bar{x}_0-\bar{x}_1}\cdot\frac{{}_{03}\sigma_1}{{}_{03}\sigma_2}$. This then is the approximate correlation our data would give if there were no selection. The value of this for each case is given in the last column but one in the table in the preceding paragraph. Its magnitude is small, and is always less than the corresponding probable error. It must be remembered, too, that there is an opposite correlation unaccounted for, since our correction for environment is not complete (see § VIII). On the whole, we believe that no error will be introduced if we assume these opposite tendencies counterbalance one another.

(XXV) If selection were perfect and its effects completed within the periods with which we have dealt in any particular case, the corresponding negative value of ${}_{03}r_{12}$ should have approached unity. On the other hand if all deaths were non-selective and purely random, a correlation not far removed from zero (see last paragraph) would have been reached. We are therefore dealing with a mixture of correlated and practically uncorrelated material, and if the conditions on which Pearson's theorem (§ III) is based were satisfied, we could deduce that the proportion of deaths which are due to selective action in each case was equal to the magnitude of the corresponding ${}_{03}r_{12}$. But there are considerable difficulties in the application of that theorem to the data of this memoir, and for the present we shall refrain from attempting to express in the form of percentages the number of deaths which are due to selection.

Conclusions.

(XXVI) The investigations of this memoir have been long and laborious, and the difficulties presented by the data have been great. Still, the general result cannot be questioned. *Natural Selection, in the form of a selective death-rate, is strongly operative in man in the early years of life.* Those data which we believe to be the best among those we have used—the Prussian figures—show very high negative correlation between the deaths in the first two years of life and those in the next eight, when allowance is made for differences in environment. We assert with great confidence that a high mortality in infancy (the first two years of life) is followed by a corresponding low mortality in childhood, and conversely. The English figures do not allow such a comprehensive survey to be undertaken, but, so far as they go they point in the same direction as the Prussian ones. The migratory tendencies in urban districts militate against the detection of selective influences there by the method of this memoir, but we express the belief that those influences are just as prevalent in industrial as in rural communities, and could be measured by other means if the data were forthcoming.

Our investigation substantiates for a general population the results found by Pearson and Ploetz for more restricted populations, and disagrees with many statements (founded, for reasons we have previously given, on inadequate evidence) made in the Local Government Board's report on Infant Mortality. It is with great reluctance that we point out this disagreement, and assert a doctrine which, in the present sentiment of society, is bound to be unpopular. We have no feelings of antagonism towards the efforts which have been made in recent years to save infant life, but we think that the probable consequences of such actions, so far as past experience can indicate them, should be completely understood. All attempts at the reduction of mortality of infancy and childhood should be made in the full knowledge of the facts of heredity. Everybody knows the extreme differences in constitutional fitness which exist in men and women. Few intelligent people can be ignorant of the fact that this constitutional fitness is inherited according to laws which are fairly definitely known. At the same time marriage is just as prevalent among those of weak stocks as among those of the vigorous, while the fertility of the former is certainly not less than that of the latter. Thus a proportion of the infants born every year must inevitably belong to the class referred to in the report as "weaklings," and, with Pearson's results before us, we are quite convinced that true infantile mortality (as distinct from the mortality due to accident, neglect, etc.—no small proportion of the whole) finds most victims from among this class. Incidentally we would here suggest that no investigation into the causes of infant and child mortality is complete until particulars are gathered by the medical officers of the constitutional tendencies and physical characters of the parents.

Our work has led us to the conclusion that infant mortality *does* effect a "weeding out" of the unfit; but, though we would give this conclusion all due emphasis, we do not wish to assert that any effort, however small, to the end of reducing this mortality is undesirable. Nobody would suggest that the difference between the infant death-rates in Oxfordshire and Glamorganshire (73 and 154 per 1000 births respectively, in 1908) was wholly due to the constitutional superiority of the inhabitants of the former county. The "weeding-out" process is not uniform. In the mining districts of South Wales, accident, negligence, ignorance and insanitary surroundings account for much. By causing improvements under these heads it may be possible to reduce the infant mortality of Glamorganshire by the survival of many who are not more unfit than are those who survive in Oxfordshire, and the social instincts of the community insist that this should be done. There are, no doubt, unpublished figures in our national archives which would enable a statistician to ascertain if the reduction in the infant mortality which has occurred since the beginning of this century has been followed by any real increase in the mortality of childhood (an apparent increase or decrease may, of course, be attributable solely to environmental influences). The correlations worked out from not very satisfactory material for the 34 districts in § IX and § X seem to suggest that the reduction *has* been inimical to the health of survivors, but we would lay no stress on a pair of correlations founded on that data. We should need to have statistics for the present century for England and Wales as complete and homogeneous as those of Prussia used in §§ XX—XXII before we could speak with any degree of certainty. Investigations on such data at intervals of ten years during which infant mortality had been almost continuously falling would be invaluable. The investigations at the beginning and end of that interval might both show a high negative correlation (corrected for environment) between the deaths in the first two years and those in the next eight, but if, at the same time, the mean number of deaths in *both* the periods were significantly reduced at the second investigation, the effect of the diminution in infantile mortality would not be detrimental so far as the first ten years were concerned. At present, however, with the results derived from the study of the Prussian data—from which we found that those provinces which from random causes (correction having been made for environment) had an excess of 100 survivors above the average at the end of the first two years of life had about 90 of this excess killed off in the next eight—before us, we feel compelled to the conclusion that, whether a reduction in the absolute mortality of infancy is or is not followed by an absolute reduction in that of childhood, the *relative* positions of different districts as regards infant mortality is, *on the whole and when allowance is made for environment*, reversed as regards the mortality of childhood for the same cohort.

From the point of view of biological science we think the results of this memoir are particularly important. As before stated, Pearson's work seems to have been ignored by many scientists, who believe that the evolution of man by means of natural selection has come to a standstill. The problem for the future is not *does* natural selection operate in man, but *how* does it operate? What particular constitutional traits or physical variations are most likely to be suppressed through mortality in early life? We have hopes that some answer to this may be rendered possible by aid of the results accruing from the medical inspection of school children.

In concluding this memoir I must express my indebtedness and gratitude to Prof. Pearson for advice and criticism during the early part of the work, for ideas derived from the records of his past labours and for his trouble in reading the whole of this memoir when in manuscript.

APPENDIX.—TABLES OF DATA.

TABLE I. *For English Registration Counties in 1903.*

County	Births in 1903	Deaths under 1 year in 1903 (see § VI)	Deaths in the first 3 years of life of those born in 1903 (see § IX)	Deaths in the first 4 years of life of those born in 1903 (see § X)	Deaths in the 4th and 5th years of life of those born in 1903 (see § IX)	Deaths between 4 and 5 years in 1907 (see § VI)	"Remaining deaths" (see § IX) in 5 years 1903–1907	Total deaths under 5 years in 1890–1900 (see § VI)	Mean Population under 5 years in 1890–1900 (see § VI)
London	131278	17223	24234	25328	1867	773	100380	335409	502241
Surrey	18585	1799	2399	2516	189	72	10437	27515	68435
Kent	24650	2712	3499	3612	243	130	14441	42676	98161
Sussex	13581	1294	1687	1766	134	55	7039	24142	60234
Hampshire	19915	1988	2696	2804	216	108	11684	34708	79953
Berkshire	6954	656	863	903	76	36	3631	11742	31135
Middlesex	25855	2921	3995	4200	357	152	17523	39652	83718
Hertfordshire	6005	505	669	688	43	24	3055	9124	25540
Buckinghamshire	4650	432	555	575	45	25	2254	7470	19847
Oxfordshire	4263	360	459	477	39	21	2172	7917	21177
Northamptonshire	9118	1041	1325	1367	74	32	5119	18105	39495
Huntingdonshire and Bedfordshire	5220	518	698	746	84	36	3005	9982	24531
Cambridgeshire	5022	508	675	702	47	20	2768	8985	22444
Essex	34218	3930	5571	5825	469	215	23510	57421	116482
Suffolk	9385	998	1300	1345	79	34	5155	17058	42472
Norfolk	11949	1373	1823	1900	129	52	7642	25153	53856
Wiltshire	6586	559	726	760	58	24	3254	10185	30193
Dorsetshire	4773	433	571	597	46	20	2267	7324	21446
Devonshire	15407	1694	2384	2485	177	76	9883	32241	69041
Cornwall	7221	711	974	1003	53	24	4071	16609	34435
Somersetshire	11004	970	1278	1325	93	46	5511	22715	57730
Gloucestershire	17018	1748	2396	2504	193	85	10160	30460	64587
Herefordshire and Shropshire	9544	970	1264	1316	83	31	4960	15586	41778
Staffordshire	41900	5986	8393	8705	583	271	34754	103270	154228
Worcestershire	14335	1774	2336	2429	158	65	9125	26659	53224
Warwickshire	28168	4055	5657	5872	401	186	23670	65704	103010
Leicestershire and Rutlandshire	13085	1741	2243	2368	191	66	9091	30074	51794
Lincolnshire	13092	1661	2176	2251	148	73	9185	25575	54993
Nottinghamshire	19728	2973	3908	4031	275	152	16100	40220	67935
Derbyshire	15127	1962	2605	2689	169	85	10488	30755	57117
Cheshire	22004	2940	4171	4358	296	109	16361	50364	86946
Lancashire	133501	20552	30151	31423	2224	952	122106	366236	498742
West Riding of Yorkshire	81140	12206	17004	17784	1299	519	66941	196863	305196
East Riding „	13552	2020	2676	2801	220	95	10707	30557	51517
North Riding „	11269	1618	2249	2350	189	88	9365	23403	44470
Durham	43819	6842	9545	9897	641	289	39177	98765	152103
Northumberland	20129	3056	4207	4372	310	145	16525	41853	69371
Cumberland and Westmorland	9080	1129	1495	1553	107	49	5943	17802	39851
Monmouthshire	11486	1502	2049	2126	143	66	9230	23138	38659
Glamorganshire	31968	4590	6241	6468	436	209	27499	71737	102096
South Wales without Glamorgan	9766	1182	1511	1556	85	40	5992	18425	40982
North Wales	12921	1586	2078	2162	129	45	8496	23324	50714

TABLE II. *For English "Areas" in* 1870.

Registration Districts comprising Area (numbered from the 1870 *Annual Report*)	Classification, A, B or C	Character of Area, U = Urban, R = Rural	Males				Females			
			Births in 1870	Deaths of those born in 1870 in the first 3 years of life	Deaths of those born in 1870 in the 4th and 5th years of life	Remaining deaths under 5 years of age from 1870–1874	Births in 1870	Deaths of those born in 1870 in the first 3 years of life	Deaths of those born in 1870 in the 4th and 5th years of life	Remaining deaths under 5 years of age from 1870–1874
1	C	U	4530	1238	105	5311	4214	961	92	4505
9	B	U	3887	994	83	4629	3762	837	102	3778
10	B	U	3917	912	76	4310	3757	843	74	3907
14	B	U	3170	804	92	3937	3068	755	72	3406
20	A	U	2676	764	65	3110	2584	590	62	2701
21	A	U	2575	684	77	3120	2573	641	77	2844
22—24 a	B	U	3317	1035	133	4645	3241	915	132	4021
24 b—25	C	U	4327	1136	123	5062	3979	909	127	4471
26	B	U	3402	938	89	4217	3237	851	95	3726
27	A	U	2497	642	75	2918	2342	555	56	2590
31	B	U	3895	957	98	4418	3865	815	85	3954
32—33	C	U	4061	1037	85	4717	4017	905	89	4174
35—36 b	C	U	4103	988	73	4158	3720	803	84	3631
37—45	B	R	3203	599	45	2370	3188	519	32	2006
46—48	A	...	2754	565	54	2384	2663	462	55	2103
49—54	B	R	3187	638	59	2620	2997	570	56	2241
55—63	B	R	3537	667	44	2662	3339	565	53	2312
64—74	C	R	4184	930	67	3666	3843	691	64	3029
75—84	B	R	3147	562	57	2315	3014	448	46	1960
85—94	B	R	3407	700	54	2941	3478	601	61	2468
95—98	A	...	2704	602	50	2673	2688	539	48	2258
99—106	B	R	3291	559	53	2457	3031	454	47	2054
107—119	A	R	2536	426	34	1673	2449	359	23	1414
120—131	B	R	3634	749	42	2948	3593	640	47	2435
132—137	C	...	4320	946	81	4289	4210	800	81	3585
138—147	B	R	3204	703	40	2688	3053	516	36	2158
148—154	A	R	2613	588	23	2363	2535	495	30	1890
155—163	A	R	2889	633	51	2585	2873	497	42	2156
164—171	A	...	2808	729	25	3002	2805	613	39	2618
172—178	A	...	2539	517	40	2178	2503	465	54	1806
179—184	A	...	2685	641	47	2501	2574	553	46	2117
185—193	B	R	3305	707	39	2671	3038	560	24	2330
194—200	B	...	3830	897	75	3974	3723	748	95	3352
201—210	B	R	3514	701	47	2768	3391	562	53	2247
211—220	A	R	2873	536	41	2126	2783	429	24	1812
221—227	A	...	2894	510	47	2321	2643	448	38	1888
228—236	B	...	3601	899	90	3397	3353	732	64	2963
237—249	B	R	3511	693	48	2759	3298	590	44	2288
250—267	B	R	3990	716	60	3000	3764	566	43	2489
268—278	A	R	2922	506	48	2160	2752	425	58	1838
279—283	A	...	2958	623	52	2542	2883	504	40	2095
284—289	B	...	3431	781	52	3420	3254	686	53	2871
290—298	B	R	3067	581	46	2106	2923	468	44	1880
299—306	A	R	2434	437	29	2011	2399	391	27	1572
307—312	B	R	3360	854	56	3380	3121	745	59	3043
313—318	A	R	2426	371	42	1827	2321	314	41	1438
319—324	A	R	2403	461	32	1847	2295	351	33	1528
325—328	A	...	2685	650	45	2665	2567	496	48	2294
329—333	C	...	4086	1021	77	4442	3811	893	59	3751
334—345	B	R	3900	794	32	3200	3866	657	41	2579
358—366	B	R	3342	683	58	2674	3145	560	58	2260
367—370	A	U	2998	827	56	3595	2986	715	64	3037
371—373	A	U	2946	853	59	3496	2845	692	56	3027
374—378	A	...	2574	506	44	2385	2447	421	55	2028
379	A	U	2772	761	67	3309	2681	659	69	2835
380—381	B	U	3632	1042	126	4832	3596	905	124	4134

TABLE II. (*continued*).

Registration Districts comprising Area (numbered from the 1870 *Annual Report*)	Classification, A, B or C	Character of Area, U = Urban, R = Rural	Males				Females			
			Births in 1870	Deaths of those born in 1870 in the first 3 years of life	Deaths of those born in 1870 in the 4th and 5th years of life	Remaining deaths under 5 years of age from 1870–1874	Births in 1870	Deaths of those born in 1870 in the first 3 years of life	Deaths of those born in 1870 in the 4th and 5th years of life	Remaining deaths under 5 years of age from 1870–1874
382	B	U	3022	936	82	4080	3034	843	83	3547
383—387	B	...	3078	741	68	3177	2948	657	61	2682
388—393	A	...	2724	550	65	2358	2718	450	44	1981
394	C	U	4382	1342	177	6012	4291	1205	174	5398
395—398	B	U	3531	890	116	4223	3354	753	117	3718
399—407	B	...	3430	748	83	3193	3260	602	79	2718
408—416	A	R	2763	620	52	2722	2702	502	48	2205
417—420	A	...	2649	793	40	3389	2420	646	36	3023
421—428	B	...	3837	811	103	3318	3699	639	81	2795
429—434	B	U	3437	632	43	2866	3247	495	54	2337
435—438	B	U	3288	761	81	3569	3114	635	71	2929
439—443	B	U	3063	782	70	3670	2999	690	72	2955
444—447	B	...	3175	656	47	3006	2929	570	63	2350
448—451	A	...	2893	643	72	3149	2835	586	77	2617
452—454	B	U	3505	939	69	4071	3256	815	63	3451
455—458	A	U	2998	728	77	2862	2893	569	62	2535
461	C	U	4484	1817	161	7483	4387	1668	196	6848
463—464	A	U	2989	712	98	3421	2920	642	102	2921
465—466	B	U	3618	1136	125	4997	3396	1024	125	4218
467—468	C	U	4127	1205	144	5255	3925	1042	127	4461
469—470	A	U	2839	786	79	3512	2808	664	71	2924
471	C	U	4050	1186	115	5236	4012	1086	124	4609
472	A	U	2580	852	93	3960	2454	756	83	3429
473	C	U	4855	1750	207	7497	4729	1533	182	6642
474	A	U	2346	665	68	3324	2303	596	60	2852
475	A	U	2367	804	71	3507	2335	685	60	3114
476—478	C	U	5064	1506	114	6548	4880	1323	119	5575
479—481	C	U	4200	1293	155	5508	3848	1042	142	4884
482—486	C	...	4669	1334	191	6432	4474	1095	196	5406
487—496	C	R	4346	999	99	4582	4177	892	84	3911
497	A	U	2612	663	60	2839	2507	510	63	2504
498	A	U	2782	752	84	3530	2645	632	79	2905
499	C	U	4937	1462	114	6805	4784	1306	131	5964
500 a, b & c	A	U	2440	734	91	3492	2288	648	90	2982
502	A	U	2471	788	88	3625	2412	589	80	3087
503—505	B	U	3525	967	113	4382	3367	792	136	3831
506—507	A	U	2620	675	86	3186	2515	553	77	2609
508	B	U	3444	1111	126	5203	3389	987	131	4560
509—514	B	...	3289	881	68	3616	3183	673	83	3037
515—519	B	...	3217	801	53	3347	3125	636	58	2914
520—526	B	...	3179	773	64	3331	3072	615	60	2754
527—539	B	R	3146	623	55	2915	2906	471	56	2518
540—541 a	A	U	2954	780	100	4085	2914	693	109	3560
541 b—544	B	U	3023	723	86	3991	2974	648	91	3422
545—547	B	U	3243	855	115	4666	2972	752	104	3819
548—549	A	U	2881	803	71	4208	2785	657	70	3511
550—551	B	U	3341	963	145	4898	3040	824	141	4306
552	A	U	2637	787	126	3745	2543	638	118	3225
553—557	A	...	2977	694	88	3324	2864	548	81	2869
558—568	B	U	3075	610	87	2924	2937	495	88	2520
569—575	B	R	3206	660	79	3176	3143	595	75	2721
576—580	B	U	3849	898	108	4342	3834	776	118	3875
581 a, b	A	U	2545	597	95	3169	2333	550	76	2730
582—583	A	U	2615	636	103	3401	2571	611	93	3038
584—585 b	A	U	2726	593	111	3081	2704	513	91	2791
586—592	B	...	3158	553	80	2523	2882	423	64	2160
614—623	B	R	3478	621	84	2857	3276	498	83	2397

Table III. *Giving the population used in Group B, § XIII.*

Registration Districts forming Area	Males	Females	Registration Districts forming Area	Males	Females
9	63974	62767	382	49956	48163
10	64557	63837	383—387	50600	51067
14	49760	49207	395—398	59216	59338
22—24 a	50844	50798	399—407	58080	57685
26	53708	52945	421—428	64599	63964
31	65102	64451	429—434	56604	55605
37—45	53121	53948	435—438	52915	52533
49—54	52406	51819	439—443	49590	49231
55—63	59892	59145	444—447	51767	51786
75—84	53742	53318	452—454	56977	57386
85—94	59584	58821	465—466	58583	57947
99—106	54289	53082	503—505	56620	56098
120—131	63280	62501	508	52768	52839
138—147	55095	53558	509—514	54173	53574
185—193	54159	54723	515—519	52415	52393
194—200	65240	63860	520—526	51234	50858
201—210	60803	60281	527—539	52984	50973
228—236	56489	55472	541 b—544	51781	50133
237—249	58561	57671	545—547	54547	52941
250—267	67883	66657	550—551	53292	52770
284—289	57991	56798	558—568	54651	52566
290—298	54103	52509	569—575	57933	56379
307—312	54667	52727	576—580	68926	68340
334—345	68171	66516	586—592	55191	52891
358—366	55433	54516	614—623	60451	58157
380—381	60825	62171			

TABLE IV. *For those born in English Rural Districts in* 1871 *and* 1872.

Registration Districts comprising Area (numbered from the 1870 *Annual Report*)	Males				Females				Males				Females			
	Births in 1871	Deaths of those born in 1871 in the first 3 years of life	Deaths of those born in 1871 in the 4th and 5th years of life	Remaining deaths under 5 years of age from 1871–1875	Births in 1871	Deaths of those born in 1871 in the first 3 years of life	Deaths of those born in 1871 in the 4th and 5th years of life	Remaining deaths under 5 years of age from 1871–1875	Births in 1872	Deaths of those born in 1872 in the first 3 years of life	Deaths of those born in 1872 in the 4th and 5th years of life	Remaining deaths under 5 years of age from 1872–1876	Births in 1872	Deaths of those born in 1872 in the first 3 years of life	Deaths of those born in 1872 in the 4th and 5th years of life	Remaining deaths under 5 years of age from 1872–1876
37—45	3232	553	54	2272	3198	414	54	1828	3260	514	49	2286	3208	414	51	1845
49—54	3139	620	64	2689	3022	492	78	2246	3144	566	92	2766	2999	437	66	2392
55—63	3234	576	59	2651	3199	466	65	2240	3516	580	55	2663	3340	504	64	2196
64—74	4070	841	76	3544	3991	674	61	2794	4225	807	91	3411	3905	641	102	2658
75—84	3094	527	59	2283	3087	444	71	1860	3102	452	42	2295	3113	423	54	1800
85—94	3508	657	65	2829	3335	559	66	2389	3587	661	72	2907	3544	548	67	2443
99—106	3150	542	56	2423	2990	470	52	2058	3359	517	72	2426	3113	449	67	2055
107—119	2445	346	33	1654	2389	275	35	1325	2652	361	23	1642	2471	288	33	1295
120—131	3889	672	61	2878	3568	527	63	2425	3780	589	57	2916	3724	528	62	2334
138—147	3266	660	52	2634	3034	498	51	2099	3456	639	49	2488	3102	459	56	2036
148—154	2753	563	50	2232	2589	405	39	1835	2686	526	32	2129	2665	376	40	1802
155—163	2869	576	52	2515	2746	451	56	2142	2902	527	40	2463	2863	448	44	2080
185—193	3222	617	51	2650	3286	532	43	2331	3270	616	54	2596	3236	542	57	2192
201—210	3601	600	44	2650	3426	493	46	2226	3644	593	54	2601	3526	504	60	2151
211—220	2869	505	50	2102	2827	424	32	1806	3016	477	30	1996	2894	403	34	1720
237—249	3564	722	54	2614	3329	531	55	2225	3302	576	45	2596	3357	514	38	2089
250—267	3894	597	66	2952	3813	490	67	2371	4054	648	71	2850	3785	508	68	2285
268—278	2904	465	67	2166	2843	391	63	1804	2899	468	55	2069	2896	397	51	1756
290—298	3092	489	52	2117	2891	393	38	1815	3137	448	51	2049	2941	384	56	1687
299—306	2358	434	34	2009	2358	375	46	1582	2486	466	39	2020	2340	372	47	1633
307—312	3271	761	78	3191	3115	639	59	2857	3344	774	79	3175	3156	668	75	2768
313—318	2449	380	72	1756	2300	341	57	1406	2526	387	56	1787	2420	300	52	1418
319—324	2342	434	42	1751	2363	326	31	1436	2382	438	46	1690	2246	356	45	1327
334—345	4025	728	62	2988	3918	608	62	2495	4041	656	96	3076	3788	530	79	2548
358—366	3258	601	62	2656	3147	543	54	2146	3233	608	54	2723	3126	493	39	2260
408—416	2778	594	65	2828	2693	480	62	2332	2718	640	90	2895	2742	493	73	2430
487—496	4426	1060	81	4527	4268	821	93	3993	4636	1056	103	4486	4403	909	92	3929
527—539	3213	656	75	3162	3034	571	108	2710	3295	680	97	3525	3108	559	102	3062
569—575	3413	656	87	3370	3310	559	102	2933	3482	705	85	3410	3280	601	75	2918
614—623	3475	638	90	2966	3342	543	71	2472	3599	654	85	3058	3219	565	89	2548

TABLE V. *For those born in Prussian Rural Districts in* 1881.

Province	Males							Females							Province
	Births	Deaths in the *first 2 years* of life	Deaths in the *first 5 years* of life	Deaths in the *second 5 years* of life	Deaths in the 15 years of life, age 5–20	Other deaths, age 0–10, in the 10 years 1881–1891 (in tens)	Other deaths, age 0–20, in the 20 years 1881–1901 (in tens)	Births	Deaths in the *first 2 years* of life	Deaths in the *first 5 years* of life	Deaths in the *second 5 years* of life	Deaths in the 15 years of life, age 5–20	Other deaths, age 0–10, in the 10 years 1881–1891 (in tens)	Other deaths, age 0–20, in the 20 years 1881–1901 (in tens)	
Königsburg	17149	4619	6213	596	872	6301	12450	16359	4076	5554	583	843	5605	11081	Königsburg
Gambinnen	13631	4081	4990	438	667	5203	10841	12888	3598	4548	429	681	4628	9656	Gambinnen
Danzig	8424	2271	2843	262	403	2926	6335	7893	1972	2528	249	361	2654	5647	Danzig
Marienwerder	15219	3918	5161	372	579	5289	11027	14389	3429	4592	393	632	4665	9679	Marienwerder
Frankfurt	13618	3313	3977	382	629	4155	8763	12812	2832	3408	378	601	3569	7488	Frankfurt
Stettin	9353	2293	2801	237	400	2757	6048	8844	1928	2452	264	425	2448	5340	Stettin
Köslin	8726	1561	2166	353	494	2336	4674	8352	1393	1938	332	470	2082	4155	Köslin
Stralsund	2340	408	499	70	121	613	1463	2161	398	494	61	111	531	1281	Stralsund
Posen	17367	4755	5671	380	638	5709	11875	16560	3798	4668	372	618	4982	10335	Posen
Bromberg	10203	2602	3291	266	408	3509	7379	9480	2271	2961	283	427	3101	6469	Bromberg
Breslau	20582	6691	7566	337	625	7596	16404	19767	5541	6365	370	740	6452	14073	Breslau
Liegnitz	13940	4374	4835	274	458	4785	10147	13223	3770	4242	258	480	4090	8694	Liegnitz
Magdeburg	10364	2594	3125	274	457	3172	6874	9690	2152	2675	307	498	2724	5961	Magdeburg
Merseburg	12873	3416	4075	323	484	4112	8953	11975	2755	3425	297	462	3560	7728	Merseburg
Erfurt	4745	1098	1363	128	200	1339	2715	4285	895	1146	99	193	1147	2376	Erfurt
Schleswig	12167	1967	2303	354	601	2474	5905	11575	1662	1992	320	617	2152	5167	Schleswig
Hanover	4628	826	1030	126	239	1122	2492	4291	645	851	108	232	962	2189	Hanover
Hildesheim	5058	972	1228	173	294	1383	3012	4814	874	1106	136	263	1192	2644	Hildesheim
Lüneburg	4893	848	1015	120	214	1081	2403	4538	700	857	130	245	920	2086	Lüneburg
Stade	4523	789	961	104	212	994	2269	4291	643	822	104	196	839	1985	Stade
Osnabrück	3723	664	796	81	164	751	1717	3434	497	612	68	158	648	1532	Osnabrück
Aurich	2649	407	469	47	100	456	1069	2569	323	390	53	111	415	949	Aurich
Münster	6051	1044	1275	133	316	1437	3873	5689	784	1013	130	352	1243	3222	Münster
Minden	7119	1296	1599	129	312	1670	3757	6766	1097	1389	166	354	1459	3393	Minden
Kassel	10898	2176	2699	286	490	2636	5377	10163	1882	2350	342	555	2348	4748	Kassel
Wiesbaden	7721	1491	1824	180	348	1737	3895	7243	1292	1589	179	384	1515	3451	Wiesbaden
Koblenz	9055	1936	2331	186	410	2278	5059	8642	1622	1987	183	398	1988	4462	Koblenz
Köln	8503	2105	2407	206	373	2470	5427	8002	1728	2042	182	382	2155	4807	Köln
Trier	11483	2233	2720	240	464	2851	6756	10683	1785	2302	285	576	2446	5952	Trier
Aachen	6291	1369	1599	115	238	1672	3892	5879	1105	1317	102	236	1419	3352	Aachen

TABLE VI. *For those born in Prussian Rural Districts in* 1882.

Province	Males							Females							Province
	Births	Deaths in the *first 2 years* of life	Deaths in the *first 5 years* of life	Deaths in the *second 5 years* of life	Deaths in the 15 years of life, age 5–20	Other deaths, age 0–10, in the 10 years 1882–1892 (in tens)	Other deaths, age 0–20, in the 20 years 1882–1902 (in tens)	Births	Deaths in the *first 2 years* of life	Deaths in the *first 5 years* of life	Deaths in the *second 5 years* of life	Deaths in the 15 years of life, age 5–20	Other deaths, age 0–10, in the 10 years 1882–1892 (in tens)	Other deaths, age 0–20, in the 20 years 1882–1902 (in tens)	
Königsburg......	18102	5192	6890	456	727	6245	12395	17481	4647	6272	461	744	5535	11018	Königsburg
Gambinnen......	14174	4358	5411	377	637	5196	10828	13701	3802	4891	368	637	4615	9638	Gambinnen
Danzig............	8657	2577	3142	213	360	2864	6322	7947	2147	2726	202	342	2578	5630	Danzig
Marienwerder...	15232	4848	6121	290	494	5289	11124	14424	3993	5222	290	511	4649	9745	Marienwerder
Frankfurt	13260	3609	4265	337	553	4183	8728	12324	2879	3526	285	507	3584	7441	Frankfurt
Stettin............	8834	2318	2840	245	411	2747	6005	8465	1953	2423	222	404	2429	5296	Stettin
Köslin	8752	1788	2589	291	451	2322	4652	8173	1506	2259	282	425	2061	4131	Köslin
Stralsund	2252	527	638	58	131	612	1480	2115	424	548	55	116	532	1295	Stralsund
Posen	17827	5186	6219	380	608	5601	11814	17088	4115	5100	400	653	4856	10226	Posen
Bromberg	10172	2996	3740	207	351	3501	7419	9680	2589	3321	220	375	3080	6477	Bromberg
Breslau	20491	7000	7889	376	711	7619	16358	19648	5768	6660	415	755	6468	14024	Breslau
Liegnitz	13648	4645	5156	253	467	4796	10072	12771	3936	4412	241	435	4093	8625	Liegnitz
Magdeburg	9846	2623	3173	251	463	3150	6824	9281	2195	2694	253	507	2695	5913	Magdeburg
Merseburg	12170	3533	4228	249	443	4135	8937	12011	2934	3594	232	424	3594	7717	Merseburg
Erfurt	4519	1084	1371	95	183	1301	2680	4172	888	1148	114	192	1116	2337	Erfurt
Schleswig	12043	1994	2352	331	592	2471	5863	11283	1584	1949	334	636	2138	5117	Schleswig
Hanover	4420	903	1070	139	249	1133	2500	4254	745	921	127	275	972	2194	Hanover
Hildesheim	4949	1028	1246	139	253	1391	2997	4801	901	1114	174	297	1193	2619	Hildesheim
Lüneburg	4718	885	1080	105	204	1084	2406	4535	668	856	134	240	917	2088	Lüneburg
Stade	4443	786	945	91	198	1015	2270	4236	694	839	125	234	852	1974	Stade
Osnabrück	3644	610	718	98	205	747	1695	3403	458	584	111	201	648	1516	Osnabrück
Aurich............	2659	382	449	45	110	448	1046	2540	284	376	51	126	448	1046	Aurich
Münster	5929	1032	1260	134	334	1471	4016	5626	875	1140	166	400	1275	3337	Münster
Minden	6889	1321	1634	130	301	1653	3744	6393	1057	1335	161	377	1437	3369	Minden
Kassel	10281	2230	2790	334	579	2574	5284	9483	1772	2284	340	556	2303	4670	Kassel
Wiesbaden	7328	1426	1724	219	418	1694	3819	6778	1104	1415	191	413	1472	3379	Wiesbaden
Koblenz	8757	1996	2341	156	354	2234	4979	8077	1673	2050	197	431	1934	4376	Koblenz
Köln	7951	2111	2444	176	364	2408	5369	7538	1700	2067	209	408	2112	4761	Köln
Trier	10921	2312	2842	243	556	2872	6733	10049	1780	2307	243	523	2466	5938	Trier
Aachen	6088	1416	1604	100	237	1672	3904	5579	1092	1322	104	234	1411	3352	Aachen

Cambridge:
PRINTED BY JOHN CLAY, M.A.
AT THE UNIVERSITY PRESS.

PUBLISHED BY THE CAMBRIDGE UNIVERSITY PRESS.

BIOMETRIKA.

A JOURNAL FOR THE STATISTICAL STUDY OF BIOLOGICAL PROBLEMS.

Founded by W. F. R. WELDON, FRANCIS GALTON and KARL PEARSON.

Edited by KARL PEARSON.

VOL. VIII., PARTS I. AND II.

I. A Third Cooperative Study of *Vespa Vulgaris.* Comparison of Queens of a Single Nest with Queens of the General Autumn Population. By E. Y. THOMSON, JULIA BELL, M.A., and KARL PEARSON, F.R.S. (With two diagrams in the text.)

II. Pigmentation of the Hair and Eyes of Children suffering from the Acute Fevers, its Effect on Susceptibility, Recuperative Power and Race Selection. By DAVID MACDONALD, M.B., Ch.B.

III. First Results from the Oxford Anthropometric Laboratory. By E. SCHUSTER, D.Sc. (With one diagram in the text.)

IV. On the Correlation between Somatic Characters and Fertility: Illustrations from the Involucral Whorl of *Hibiscus.* By J. ARTHUR HARRIS, Ph.D. (With one diagram in the text.)

V. Anthropometry of Modern Egyptians. By J. I. CRAIG, M.A., F.R.S.E. (With two diagrams in the text.)

VI. The Teacher's Estimation of the General Intelligence of School Children. By H. WAITE, M.A.

VII. On the Significance of the Teacher's Appreciation of General Intelligence. By WALTER H. GILBY, B.Sc., assisted by KARL PEARSON, F.R.S. (With one diagram in the text.)

VIII. The Danger of Certain Formulae suggested as Substitutes for the Correlation Coefficient. By DAVID HERON, D.Sc. (With seven diagrams in the text.)

IX. Cranial Type-Contours. By the late R. CREWDSON BENINGTON, M.D., prepared for press by KARL PEARSON, F.R.S. (With thirty-two plates in the text and thirty-two copies on tissue in pocket.)

X. The Opsonic Index. "Mathematical Error and Functional Error." By KARL PEARSON, F.R.S. (With nine diagrams in the text.)

Miscellanea: (i) A note on the Age Distribution of Deaths from Diabetes Mellitus. By G. D. MAYNARD. (With three diagrams in the text.) (ii) On Pulse and Breathing Rates and their Relation to Stature. By JULIA BELL, M.A. (With two diagrams in the text.) (iii) The Tooth-Brush and Decayed Teeth. By FRANK ROCK, M.D. (iv) Note on High Values of β_1, β_2. By GEORG DUNCKER, Ph.D. (v) Further Remarks on the Law of Ancestral Heredity. By KARL PEARSON, F.R.S. (vi) The Inheritance of Fertility. By R. E. LLOYD, D.Sc. (Lond.). (With one diagram in the text.) (vii) Remarks on Professor Lloyd's Note on Inheritance of Fertility. By KARL PEARSON, F.R.S. (viii) On the Probability that two Independent Distributions of Frequency are really Samples from the same Population. By KARL PEARSON, F.R.S. (ix) On a Correction needful in the Case of the Correlation Ratio. By KARL PEARSON, F.R.S.

The subscription price, payable in advance, is 30*s. net* per volume (post free); single numbers 10*s. net.* Volumes I., II., III., IV., V., VI. and VII. complete, 30*s. net* per volume. Bound in Buckram 34*s.* 6*d. net* per volume. Index to Volumes I. to V. 2*s. net.* Subscriptions may be sent to C. F. CLAY, Manager, Cambridge University Press, Fetter Lane, London, E.C., either direct or through any bookseller, and communications respecting advertisements should also be addressed to C. F. CLAY.

Till further notice, new subscribers may obtain Vols. I.—VI. together for £7. 10*s.*, or bound in Buckram for £8. 10*s.*

BIOMETRIC LABORATORY PUBLICATIONS

ISSUED BY MESSRS DULAU & CO., LTD., 37, SOHO SQUARE, LONDON, W.

Questions of the Day and of the Fray.

I. The Influence of Parental Alcoholism on the Physique and Ability of the Offspring. A Reply to the Cambridge Economists. By KARL PEARSON, F.R.S. *Issued.* Price 1*s. net.*

II. Mental Defect, Mal-Nutrition, and the Teacher's Appreciation of Intelligence. A Reply to Criticisms of the Memoir on 'The Influence of Defective Physique and Unfavourable Home Environment on the Intelligence of School Children.' By DAVID HERON, D.Sc. *Issued.* Price 1*s. net.*

III. An Attempt to correct some of the Misstatements made by Sir VICTOR HORSLEY, F.R.S., F.R.C.S., and MARY D. STURGE, M.D., in their Criticisms of the Galton Laboratory Memoir: 'A First Study of the Influence of Parental Alcoholism,' etc. By KARL PEARSON, F.R.S. *Issued.* Price 1*s. net.*

IV. The Fight against Tuberculosis and the Death-rate from Phthisis. BY KARL PEARSON, F.R.S. *Issued.* Price 1*s. net.*

UNIVERSITY OF LONDON (UNIVERSITY COLLEGE).

DRAPERS' COMPANY RESEARCH MEMOIRS

A. *Technical Series.*

I. On a Theory of the Stresses in Crane and Coupling Hooks with Experimental Comparison with Existing Theory. By E. S. Andrews, B.Sc. Eng., assisted by Karl Pearson, F.R.S. *Issued.* Price 3*s. net.*

II. On some Disregarded Points in the Stability of Masonry Dams. By L. W. Atcherley, assisted by Karl Pearson, F.R.S. *Issued.* Price 3*s.* 6*d. net.*

III. On the Graphics of Metal Arches, with special reference to the Relative Strength of Two-pivoted, Three-pivoted and Built-in Metal Arches. By L. W. Atcherley and Karl Pearson, F.R.S. *Issued.* Price 6*s. net.*

IV. On Torsional Vibrations in Axles and Shafting. By Karl Pearson, F.R.S. *Issued.* Price 6*s. net.*

V. An Experimental Study of the Stresses in Masonry Dams. By Karl Pearson, F.R.S., and A. F. Campbell Pollard, assisted by C. W. Wheen, B.Sc. Eng., and L. F. Richardson, B.A. *Issued.* Price 7*s. net.*

VI. On a Practical Theory of Elliptic and Pseudo-elliptic Arches, with special reference to the ideal Masonry Arch. By Karl Pearson, F.R.S., W. D. Reynolds, B.Sc. Eng., and W. F. Stanton, B.Sc. Eng. *Issued.* Price 4*s. net.*

B. *Biometric Series.*

I. Mathematical Contributions to the Theory of Evolution.—XIII. On the Theory of Contingency and its Relation to Association and Normal Correlation. By Karl Pearson, F.R.S. *Issued.* Price 4*s. net.*

II. Mathematical Contributions to the Theory of Evolution.—XIV. On the Theory of Skew Correlation and Non-linear Regression. By Karl Pearson, F.R.S. *Issued.* Price 5*s. net.*

III. Mathematical Contributions to the Theory of Evolution.—XV. On the Mathematical Theory of Random Migration. By Karl Pearson, F.R.S., with the assistance of John Blakeman, M.Sc. *Issued.* Price 5*s. net.*

IV. Mathematical Contributions to the Theory of Evolution.—XVI. On Further Methods of Measuring Correlation. By Karl Pearson, F.R.S. *Issued.* Price 4*s. net.*

V. Mathematical Contributions to the Theory of Evolution.—XVII. On Homotyposis in the Animal Kingdom. By Ernest Warren, D.Sc., Alice Lee, D.Sc., Edna Lea-Smith, Marion Radford, and Karl Pearson, F.R.S. *Shortly.*

VI. Albinism in Man. By Karl Pearson, E. Nettleship, and C. H. Usher. *Text,* Part I. *Atlas,* Part I. *Issued.* Price 35*s. net.*

VII. Albinism in Man. By Karl Pearson, E. Nettleship, and C. H. Usher. *Text,* Part II. *Atlas,* Part II. *Immediately.*

C. *Studies in National Deterioration.*

I. On the Relation of Fertility in Man to Social Status, and on the changes in this Relation that have taken place in the last 50 years. By David Heron, M.A. *Issued.* Price 3*s. net.*

II. A First Study of the Statistics of Pulmonary Tuberculosis: Inheritance. By Karl Pearson, F.R.S. *Issued.* Price 3*s. net.*

III. A Second Study of the Statistics of Pulmonary Tuberculosis: Marital Infection. By the late E. G. Pope. Edited and revised by Karl Pearson, F.R.S. with an Appendix on Assortive Mating from Data reduced by Ethel M. Elderton. *Issued.* Price 3*s. net.*

IV. The Health of the School-Child in relation to its Mental Characters. By Karl Pearson, F.R.S. *Shortly.*

V. On the Inheritance of the Diatheses of Phthisis and Insanity. A Statistical Study based upon the Family History of 1500 Criminals. By Charles Goring, M.D., B.Sc. *Issued.* Price 3*s. net.*

VI. A Third Study of the Statistics of Pulmonary Tuberculosis: The Mortality of the Tuberculous and Sanatorium Treatment. By W. Palin Elderton, F.I.A. and S. J. Perry, A.I.A. *Issued.* Price 3*s. net.*

VII. On the Intensity of Natural Selection in Man. By E. C. Snow, M.A. *Issued.* Price 3*s. net.*

PUBLISHED BY DULAU & CO., LTD., 37, SOHO SQUARE, LONDON, W.